COUPLING TIMES WITH AMBIGUITIES FOR PARTICLE SYSTEMS AND APPLICATIONS TO CONTEXT-DEPENDENT DNA SUBSTITUTION MODELS

JEAN BÉRARD, DIDIER PIAU

Abstract. We define a notion of coupling time with ambiguities for interacting particle systems, and show how this can be used to prove ergodicity and to bound the convergence time to equilibrium and the decay of correlations at equilibrium. A motivation is to provide simple conditions which ensure that perturbed particle systems share some properties of the underlying unperturbed system. We apply these results to context-dependent substitution models recently introduced by molecular biologists as descriptions of DNA evolution processes. These models take into account the influence of the neighboring bases on the substitution probabilities at a site of the DNA sequence, as opposed to most usual substitution models which assume that sites evolve independently of each other.

Contents

1. Introduction and motivations

This paper is devoted to interacting particle systems on the integer line $\mathbb{Z}$ with finite state space S, whose dynamics is characterized by a finite list $\mathfrak{R}$ of stochastic transition rules. We now give an informal description of the dynamics that we consider for these systems, and we postpone a proper mathematical definition to section 2.

Date: February 2, 2008.
2000 *Mathematics Subject Classification.* 60J25, 60K35, 92D20.
Key words and phrases. Interacting particle systems, Coupling, Perturbations, Stochastic models of nucleotide substitutions.

1.1. **Construction of interacting particle systems dynamics.** We begin with some vocabulary. A *state* s is an element of S, a *site* x is an element of $\mathbb{Z}$, a *configuration* $\xi := (\xi(x))_{x\in\mathbb{Z}}$ is an element of $S^{\mathbb{Z}}$. A *rule* $\mathcal{R} := (c, r)$ is based on a context c and characterized by a rate r. A *context* is a triple $c := (A, \ell, s)$, where A is a finite subset of $\mathbb{Z}$, ℓ is a subset of S^A, s is a state, and r is a *rate*, that is, a non-negative real number.

We say that a configuration ξ and a context $c = (A, \ell, s)$, or any rule $\mathcal{R} = (c, r)$ based on c, are *compatible at site* x if A is empty, or if A is not empty and $\xi(x+A)$ belongs to ℓ, where $\xi(x + A)$ is the element of S^A defined as

$$\xi(x + A) := (\xi(x + y))_{y\in A}.$$

The interacting particle system is a Markov process $(X_t)_t$ on $S^{\mathbb{Z}}$ whose dynamics is characterized by a given finite list $\mathfrak{R} := (\mathcal{R}_i)_{i\in\mathfrak{I}}$ of stochastic transition rules, as follows: for any time t, if a rule $\mathcal{R}_i = (c_i, r_i)$ in $\mathfrak{R}$ with $c_i = (A_i, \ell_i, s_i)$ is compatible with X_t at site x, then $X_{t+\mathrm{d}t}(x) = s_i$ with probability $r_i\mathrm{d}t + o(\mathrm{d}t)$, independently of every other rule in $\mathfrak{R}$, compatible with X_t at site x or elsewhere.

A classical way to give a more explicit construction of such particle systems uses the so-called graphical representation (see for instance [11] page 142 for a discussion in the context of voter models). This amounts to a stochastic flow based on Poisson processes: given a time t and an initial condition ξ in $S^{\mathbb{Z}}$ imposed at time t, the Poisson processes determine the state of the particle system at every time greater than t. Once again informally, to every site x and rule $\mathcal{R}_i = (c_i, r_i)$ in $\mathfrak{R}$ corresponds a homogenous Poisson process $\Psi(x, i)$ on the real line $\mathbb{R}$ with rate r_i, and the points of $\Psi(x, i)$ are the random times at which the rule $\mathcal{R}_i$ is applied to the state at site x. Specifically, for every rule $\mathcal{R}_i = (c_i, r_i)$ in $\mathfrak{R}$ with context $c_i = (A_i, \ell_i, s_i)$, if t belongs to $\Psi(x, i)$ and if $\mathcal{R}_i$ and X_{t^-} are compatible at site x, then $X_t(x) = s$; otherwise, $X_t(x) = X_{t-}(x)$. See section 2 for a proper definition.

1.2. **Coupling times.** Within this framework, various notions of coupling times can be defined. In this paper, an *ordinary coupling time* is an almost surely finite random variable T with negative values, measurable with respect to the family $(\Psi(x, i))_{(x,i)\in\mathbb{Z}\times\mathfrak{I}}$ of Poisson processes, and such that, for every time $u < T$, if the dynamics starts at time u, the state of site $x = 0$ at time $t = 0^-$ is the same for every initial condition at time u. This definition corresponds to a coupling from the past, as opposed to the usual notion of forward coupling.

As soon as such coupling times exist, the particle system is ergodic. Furthermore, estimates on the tail of T yield estimates on the rate of convergence to equilibrium, and additional assumptions on the coupling time yield estimates on the decay of correlations. Consider now the set of points

$$\mathcal{T} := \bigcup_{(x,i)} (\Psi(x, i) \cap [T, 0[) \times \{x\},$$

where the union runs over every x in $\mathbb{Z}$ and i in $\mathfrak{I}$. A point in $\mathcal{T}$ corresponds to a transition that may or may not be performed between the times $t = T$ and $t = 0^-$, depending on the initial condition at time $v < T$. When, for a given (u, x) in $\mathcal{T}$, there indeed exists $v < T$ and two distinct initial conditions at time v such that, for one of these initial conditions, the transition proposed by (u, x) is performed, while it is not performed when the other initial condition is used, we say that an

ambiguity arises at (u, x). By the definition of an ordinary coupling time, one sees that, for each time in $\mathcal{T}$, either there is no ambiguity associated with it, or there is an ambiguity that has no influence on the state of site $x = 0$ at time $t = 0^-$.

We can now define, once again informally, the notion of coupling time with ambiguities. This is a pair (H, T), where T is a random variable with negative values, measurable with respect to the family $(\Psi(x, i))_{(x,i)\in\mathbb{Z}\times\mathfrak{I}}$ of Poisson processes and H is a finite random subset of the set $\mathcal{T}$ defined above, enjoying the stopping property, and such that the following property holds: for any two initial conditions at time $u < T$ such that the ambiguities associated with the elements of H are resolved in the same way (that is, a transition corresponding to an element of H is either performed for both initial conditions, or not performed for both initial conditions), the state of site $x = 0$ at time $t = 0^-$ is the same for both initial conditions.

One sees that, if (H, T) is a coupling time with ambiguities, T may or may not be an ordinary coupling time. However, the only ambiguities that may prevent T from being an ordinary coupling time are those associated to the points in H. As a consequence, in the degenerate case when H is empty, T is indeed an ordinary coupling time.

Informally, our main result is that, if the random set H contains few enough points on average (we call *subcritical* any coupling time with ambiguities enjoying this property), it is possible to build an ordinary coupling time from (H, T), thus proving ergodicity of the particle system. Moreover, more specific estimates and assumptions about the set H provide estimates on this ordinary coupling time, that are suitable to study the rate of convergence to equilibrium of the particle system and the decay of its correlations.

The construction of an ordinary coupling time from a subcritical coupling time with ambiguities is described in section 4. The principle of this construction is to apply iteratively coupling times with ambiguities, looking further and further into the past, until every ambiguity is eventually resolved.

1.3. **Perturbed particle systems.** We now describe how these results allow to study some perturbed particle systems. We assume that the list of transition rules is of the form $\mathfrak{R} = (\mathfrak{R}_i)_{i\in\mathfrak{I}^o\cup\mathfrak{I}^p}$, where $\mathfrak{I}^o$ and $\mathfrak{I}^p$ are disjoint sets, the family $\mathfrak{R}^o := (\mathfrak{R}_i)_{i\in\mathfrak{I}^o}$ being the family of so-called non-perturbative rules, while $\mathfrak{R}^p := (\mathfrak{R}_i)_{i\in\mathfrak{I}^p}$ is the family of so-called perturbative rules.

We call the interacting particle system based on the whole family of rules $\mathfrak{R}$ the perturbed system and the system based on the family of non-perturbative rules $\mathfrak{R}^o$ the unperturbed system.

A general problem about perturbations of particle systems is to relate the properties of the perturbed system such as ergodicity, speed of convergence to equilibrium or decay of correlations at equilibrium, to those of the unperturbed system, when the transition rates attached to the perturbative rules are small enough. In this context, we wish to mention two results, one on the negative side and one on the positive side:

(1) Small perturbations of ergodic particle systems may not be ergodic. For a well-known example, consider the two-dimensional Ising model. Its dynamics is ergodic at the critical inverse temperature β_c and not ergodic at

any inverse temperature $\beta > \beta_c$, see [10] (page 204 and Theorem 2.16 on page 195) for instance.

(2) Small perturbations of particle systems whose coordinates evolve independently are ergodic, see [10] (Theorem 4.1 on page 31) for instance.

Depending on the assumptions one makes about the unperturbed system, and on the perturbations one considers, one can use various methods to deal with this problem. For an example of the use of L^2 techniques and spectral gap estimates, see [9]. For an example of the use of cluster-expansion estimates, see [13]. For examples involving Lyapounov function techniques, in the slightly different context of perturbations of Markov chains, see [1, 6, 15, 16].

The approach of this paper is based on coupling. The basic idea is that, in some situations, it should be possible to rely on the coupling properties of the unperturbed system to devise a coupling time with ambiguities (H, T) for the perturbed system, with the property that, when the rates associated with perturbative rules are small enough, the coupling time with ambiguities is subcritical. We do not provide an abstract formulation of this idea, but, as an illustration, we give two concrete examples in section 5, in the context of stochastic models of nucleotide substitution in molecular evolution, recently studied in [3].

1.4. **Organization of the paper.** Section 2 contains a formal definition of the interacting particle systems studied in this paper, their construction by means of Poisson processes, the definition of the notion of coupling time with ambiguities, and some notations. Section 3 contains the main results, whose proofs are in section 4. Section 5 applies these theoretical results to a concrete case, namely a class of stochastic models of context-dependent nucleotide substitution, recently introduced by molecular biologists, and whose study was our initial motivation for the results in this paper.

2. Formal setting

2.1. **Preliminary definitions and notations.** In this paper, particle systems are continuous-time Markov processes on $S^{\mathbb{Z}}$, where S denotes a finite set. Sites x are elements of $\mathbb{Z}$, states s are elements of S and configurations $\xi = (\xi(x))_{x\in\mathbb{Z}}$ are elements of $S^{\mathbb{Z}}$. The space $\mathcal{C} := \mathcal{C}([0, +\infty[, S^{\mathbb{Z}})$ is the space of càdlàg functions on $[0, +\infty[$ with values in $S^{\mathbb{Z}}$. For every nonnegative time t, $X_t : \mathcal{C} \to S^{\mathbb{Z}}$ is the canonical coordinate map on $\mathcal{C}$, hence $X_t(\omega) := \omega(t)$ for every ω in $\mathcal{C}$. The space $\mathcal{C}$ is endowed with the cylindrical σ-algebra $\sigma((X_t)_{t\geqslant 0})$. For every nonempty subset B of $\mathbb{Z}$, $\pi_B : S^{\mathbb{Z}} \to S^B$ is the canonical projection defined by

$$\pi_B(\xi) := (\xi(x))_{x\in B}.$$

For every site x, $\pi_x := \pi_{\{x\}}$. For every site y, $\vartheta_y : S^{\mathbb{Z}} \to S^{\mathbb{Z}}$ is the canonical translation of $S^{\mathbb{Z}}$ defined by

$$\vartheta_y(\xi) := (\xi(x + y))_{x\in\mathbb{Z}}.$$

For every configuration ξ, site x and state s, the configuration $\xi^{x,s}$ is defined by $\xi^{x,s}(x) := s$, and $\xi^{x,s}(y) := \xi(y)$ for every site $y \neq x$.

Finally, $\mathbb{R}_+ := [0, +\infty[$, $\mathbb{R}_- :=]-\infty, 0]$, and $C(S^{\mathbb{Z}})$ is the space of the functions F defined on $S^{\mathbb{Z}}$ such that the following series converges:

$$\sum_{x\in\mathbb{Z}} \sup\{\, |F(\xi^{x,s}) - F(\xi)| \,;\, \xi \in S^{\mathbb{Z}},\, s \in S\}.$$

2.2. **Specifications by transition rules.** Recall that one is given a finite list of transition rules

$$\mathfrak{R} = (\mathcal{R}_i)_{i\in\mathfrak{I}}, \qquad \mathcal{R}_i = (c_i, r_i),$$

indexed by a finite set $\mathfrak{I}$. For every i in $\mathfrak{I}$, the rate r_i of the rule $\mathcal{R}_i$ is a nonnegative real number and its context $c_i = (A_i, \ell_i, s_i)$ is characterized by a finite subset A_i of $\mathbb{Z}$, a subset ℓ_i of S^{A_i} and a state s_i in S.

The list of rules $\mathfrak{R}$ yields a definition of the dynamics of the process through its infinitesimal generator $\mathfrak{L}$, as follows: for every function F in $C(S^{\mathbb{Z}})$ and every configuration ξ,

$$(\mathfrak{L}F)(\xi) := \sum_{(x,i)} r_i\, \mathbf{1}\{\pi_{A_i}(\vartheta_x\xi) \in \ell_i\}\, (F(\xi^{x,s_i}) - F(\xi)),$$

where the sum enumerates every site x in $\mathbb{Z}$ and rule index i in $\mathfrak{I}$. Here and below, we adopt the convention that $\mathbf{1}\{\pi_\emptyset(\xi) \in \emptyset\} = 1$ for every ξ in $S^{\mathbb{Z}}$.

The fact that the generator $\mathfrak{L}$ indeed defines a Feller Markov process $(X_t)_{t\geqslant 0}$ on $S^{\mathbb{Z}}$ is a standard result, see [10] for instance. For every configuration ξ, $\mathbb{Q}^\xi$ denotes the probability measure on $\mathcal{C}$ which corresponds to the initial condition $X_0 = \xi$.

Distinct families of stochastic transition rules may lead to the same infinitesimal generator $\mathfrak{L}$. As a consequence, the probability measures $\mathbb{Q}^\xi$ do not uniquely determine $\mathfrak{R}$ in general, and several families of rules are compatible with the same Markov process. However, distinct families of rules do lead to distinct versions of the construction presented in the next section, so the coupling properties of this construction that are of essential use in this paper, may differ substantially from one family of rules to another, even when the corresponding infinitesimal generators are the same.

2.3. **Dynamics based on Poisson processes.** The infinitesimal generator $\mathfrak{L}$ defined above is the usual way to specify a dynamics from a finite collection of transition rules. However, the coupling times that we consider in this paper are formulated in terms of another construction, the so-called graphical construction of the dynamics, see [10], which is based on a family of Poisson processes corresponding to transition times.

We now describe this construction in details. Since we will be interested in coupling from the past, we only have to define the dynamics of the process for the "past" times $t \leqslant 0$.

2.3.1. *Poisson processes.* The individual sample space for the Poisson processes is the set

$$\Omega_0 := \{(t_k)_{k\geqslant 1} \subset \mathbb{R}\,;\, \forall k \geqslant 1,\, t_{k+1} < t_k < 0,\, \lim_{n\to+\infty} t_n = -\infty\}.$$

We equip Ω_0 with the σ-algebra $\mathcal{F}_0$ induced by the product Borel σ-algebra on the space of real valued sequences. For every $k \geqslant 1$, the coordinate map $T_k : \Omega_0 \to \mathbb{R}$ is defined by

$$T_k((t_n)_{n\geqslant 1}) := t_k.$$

For every site x in $\mathbb{Z}$ and rule index i in $\mathfrak{I}$, $\mathbb{P}_x^i$ is the probability measure on $(\Omega_0, \mathcal{F}_0)$ such that the sequence $(T_k)_{k\geqslant 1}$ is a Poisson process on $\mathbb{R}_-$ with rate r_i.

To define the dynamics of interest, we introduce a family of processes on the sample probability space

$$(\Omega, \mathcal{F}, \mathbb{P}) := \bigotimes_{(x,i)} (\Omega_0, \mathcal{F}_0, \mathbb{P}_x^i),$$

where the $\otimes$ product enumerates every site x in $\mathbb{Z}$ and every rule index i in $\mathfrak{I}$. For every x in $\mathbb{Z}$, i in $\mathfrak{I}$, $k \geqslant 1$, and ψ in Ω, one defines

$$\Psi(x,i,k)(\psi) := T_k(\psi(x,i)), \quad \Psi(x,i)(\psi) := \psi(x,i),$$

and

$$\Psi(x)(\psi) := (\psi(x,i))_{i\in\mathfrak{I}}, \quad \Psi(\psi) := \psi.$$

With these notations, $\Psi(x,i) = (\Psi(x,i,k))_{k\geqslant 1}$. In the sequel, $\Psi(x,i)$ denotes also the random set $\bigcup_{k\geqslant 1} \{\Psi(x,i,k)\}$ and the context should make clear which one of these two notations is in use. The same convention applies to $\Psi(x)$ and Ψ. Finally, $\mathcal{F}(x,i)$ is the sub-σ-algebra of $\mathcal{F}$ generated by $\Psi(x,i)$.

Remark 2.1. *In our context, it is necessary to use an indexation of the various random variables Ψ and of related quantities by rule indices i in $\mathfrak{I}$ instead of an indexation by the rules $\mathcal{R}$ in $\mathfrak{R}$ themselves. To see why, consider the case when two rules $\mathcal{R}_i$ and $\mathcal{R}_j$ with $i \neq j$ in $\mathfrak{I}$ are described by the same contexts $c_i = c_j$ and the same rates $r_i = r_j$. Then $\mathcal{R}_i = \mathcal{R}_j$ but we want to consider the addition of their two effects, which could also be described by the single rule with context c_i and rate $2r_i$.*

2.3.2. *Flows.* Let $\Omega_1 \subset \Omega$ denote the event that $\Psi(x,i,k) \neq \Psi(x',i',k')$ for every $(x,i,k) \neq (x',i',k')$. Then $\mathbb{P}(\Omega_1) = 1$, that is, almost surely, for each time $t \leqslant 0$, t belongs to exactly one set $\Psi(x,i,k)$ or t belongs to none of them.

Definition 2.2 (Direct influences)**.** *The direct influence process is the random process* Influ *defined on $\mathbb{R}_- \times \mathbb{Z}$ as follows. Let $t \leqslant 0$ and x in $\mathbb{Z}$. If t belongs to a unique set $\Psi(x,i,k)$, let*

$$\mathrm{Influ}(t) := \{t\} \times (x + A_i).$$

Otherwise, let $\mathrm{Influ}(t) := \{(t,x)\}$*. Conversely, for every site x, let*

$$\mathrm{Influ}^{-1}(t,x) := t.$$

When A_i is empty, this definition implies that $\mathrm{Influ}(t)$ is empty.

Definition 2.3 (Preceding times)**.** *For every times $u < t \leqslant 0$ and site x, the preceding time* $\mathrm{Preced}_u(t,x)$ *at x after u and before t is the random variable with values in $\mathbb{R}_- \times \mathbb{Z}$ defined by*

$$\mathrm{Preced}_u(t,x) := (\sup\,]u,t[\cap \Psi(x), x)\,,$$

with the convention that $\mathrm{Preced}_u(t,x) := (u,x)$ *if $]u,t[\cap\Psi(x)$ is empty.*

Definition 2.4 (Multilevel influences)**.** *For every site x and times $u < t \leqslant 0$, we define inductively a sequence $(\mathrm{Influ}_k(u,t,x))_{k\geqslant 0}$ of random sets, called the influences of site x after u and before t at level k, as follows.*

- *For $k = 0$, let $\mathrm{Influ}_0(u,t,x) := \mathrm{Influ}(t,x)$;*
- *For every $k \geqslant 0$,*

$$\mathrm{Influ}_{k+1}(u,t,x) := \mathrm{Influ}_k(u,t,x) \cup \mathrm{Influ}(\mathrm{Preced}_u(\mathrm{Influ}_k(u,t,x))).$$

Additionally, the complete influence of site x after u and before t is

$$\mathrm{Influ}_\infty(u,t,x) := \bigcup_{k=0}^{+\infty} \mathrm{Influ}_k(u,t,x).$$

Let $\Omega_2 \subset \Omega$ denote the event that $\mathrm{Influ}_\infty(u,t,x)$ is a finite set for every site x and couple (u,t) of times such that $u < t \leqslant 0$. Then $\mathbb{P}(\Omega_2) = 1$ (see [11] page 142, for instance).

We define a random flow Φ on $S^{\mathbb{Z}} \times \mathbb{R} \times \mathbb{R} \times \mathbb{Z}$, such that, for every $u \leqslant t \leqslant 0$, $\Phi(\xi,u,t,x)$ is the x-coordinate of the configuration at time t which one obtains by applying the moves described by Ψ to the configuration ξ at time u.

The definition of Φ is based on the following recursive procedure.

Assume first that $\Omega_1 \cap \Omega_2$ holds. Fix a site x and times $u \leqslant t \leqslant 0$. If $u = t$, let $\Phi(\xi,u,t,x) := \xi(x)$. If $u < t$, consider first the case where t is not in $\Psi(x)$. Then $\mathrm{Influ}(t,x) = \{(t,x)\}$, and we use a recursive call to the definition of Φ by letting

$$\Phi(\xi,u,t,x) := \Phi(\xi, \mathrm{Preced}_u(t,x)).$$

Otherwise, $t = \Psi(x,i,k)$ for exactly one rule index i in $\mathfrak{I}$ and one index $k \geqslant 1$. Consider then the set $\mathrm{Preced}_u(\mathrm{Influ}(t,x))$. If this set reduces to $\{(u,x)\}$, then, for any (y,t) in $\mathrm{Influ}(t,x)$, no rule applies at site y between the times u and t. Then, let $\chi(y) := \xi(y)$. Otherwise, $\mathrm{Preced}_u(\mathrm{Influ}(t,x))$ is not reduced to $\{(u,x)\}$. For every element (y,t) of $\mathrm{Influ}(t,x)$, we use a recursive call to the definition of Φ and let

$$\chi(y) := \Phi(\xi,u,\mathrm{Preced}_u(t,y)).$$

Then, if $\chi(x+A_i)$ belongs to ℓ_i (remember that this is automatically the case when A_i is empty), let $\Phi(\xi,u,t,x) := s_i$, and say that $\Psi(x,i,k)$ is performed when one starts from configuration ξ at time u. Otherwise, let $\Phi(\xi,u,t,x) := \xi(x)$ and say that $\Psi(x,i,k)$ is not performed when one starts from configuration ξ at time u. For the sake of definiteness, if $\Omega_1 \cap \Omega_2$ does not hold, let $\Phi(\xi,u,t,x) := \xi(x)$, and say that $\Psi(x,i,k)$ is not performed, whatever the value of (x,i,k) is. This ends the description of the construction of Φ.

The fact that, on $\Omega_1 \cap \Omega_2$, $\mathrm{Influ}_\infty(u,t,x)$ is a finite set, guarantees that the above procedure involves only a finite number of recursive calls to the definition of Φ and leads to a consistent definition of Φ. Moreover, one can check that the fact that $\Psi(x,i,k)$ is performed or not does not depend on the value of t, but only on ξ, u, and, of course, Ψ. The proof of the proposition below is standard.

Proposition 2.5 (Flow properties)**.** *The flow Φ enjoys the following properties.*

- *For every times $u < v < t \leqslant 0$ and site x,*

$$\Phi(\Phi(\xi,u,v,\cdot),v,t,x) = \Phi(\xi,u,t,x).$$

- *For every time $u \leqslant 0$, the distribution of the càdlàg random process*

$$(\Phi(\xi, u, u+t, \cdot))_{0 \leqslant t \leqslant -u}$$

is the distribution of $(X_t)_{0 \leqslant t \leqslant -u}$ with respect to $\mathbb{Q}^\xi$.

A motivation to give the details of the construction of Φ was to be able to define the following random variable.

Definition 2.6 (Performance indicator)**.** *The performance indicator of rank $k \geqslant 1$ for the rule index i at site x, starting from configuration ξ at time u, is*

$$\mathrm{Perf}(\xi, u, x, i, k) := \mathbf{1}\{\Psi(x, i, k) \text{ is performed when starting from } \xi \text{ at time } u\}.$$

2.3.3. *Measurability and shifts.* For every time $t \leqslant 0$ and rule index i in $\mathfrak{I}$, let $K_t(x, i)$ denote the random variable on $(\Omega, \mathcal{F}, \mathbb{P})$ defined by

$$K_t(x, i) := \max\{k \geqslant 1\,;\, \Psi(x, i, k) \geqslant t\} \cup \{0\}.$$

For every time $t \leqslant 0$ and site y, the space-time-shift $\sigma_{t,y}$ is defined on Ω by

$$[\sigma_{t,y}(\Psi)](x, i, k) := \Psi\left(x + y, i, k + K_t(x + y, i)\right) - t.$$

Then $\mathbb{P}$ is invariant with respect to every $\sigma_{t,y}$ and

$$\mathbf{1}_{\Omega_1 \cap \Omega_2} \leqslant \mathbf{1}_{\Omega_1 \cap \Omega_2} \circ \sigma_{t,y}.$$

The behavior of the flow under the action of the shift is described by our next lemma, whose proof is left to the reader.

Lemma 2.7. *On $\Omega_1 \cap \Omega_2$, for every sites x and y and times $u \leqslant v \leqslant 0$ and $t \leqslant 0$,*

$$\Phi(\xi, u, v, x) \circ \sigma_{t,y} = \Phi(\xi, u + t, v + t, x + y),$$

and, for every rule index i in $\mathfrak{I}$ and index $k \geqslant 1$,

$$\mathrm{Perf}(\xi, u, x, i, k - K_t(x, i)) \circ \sigma_{t,y} = \mathrm{Perf}(\xi, u + t, x + y, i, k).$$

Let $\Gamma := \bigcup_{n \geqslant 0} \{n\} \times \mathbb{R}^n$ and $\mathcal{F}^\Gamma$ the σ-algebra on Γ generated by the sets $\{n\} \times B$, for every $n \geqslant 0$ and every Borel subset B of $\mathbb{R}^n$.

Definition 2.8. *For every time $t \leqslant 0$, let $\mathcal{F}^+(t)$ denote the sub-σ-algebra of $\mathcal{F}$ generated by the family of maps $\theta_t^+(x, i) : \Omega \to \Gamma$, for every site x in $\mathbb{Z}$ and rule index i in $\mathfrak{I}$, defined by*

$$\theta_t^+(x, i) := \left(K_t(x, i)\,;\, \{\Psi(x, i, k)\,;\, 1 \leqslant k \leqslant K_t(x, i)\}\right).$$

More generally, if U is a random variable defined on $(\Omega, \mathcal{F})$ with values in $\mathbb{R}_-$, $\mathcal{F}^+(U)$ denotes the sub-σ-algebra of $\mathcal{F}$ generated by the maps $\theta_U^+(x, i)$.

One can view $\mathcal{F}^+(t)$ as the σ-algebra of the events posterior to the time t.

2.4. **Coupling times with ambiguities.**

Definition 2.9. *Let $H = \{H(x,i,k)\,;\, x \in \mathbb{Z},\, i \in \mathfrak{I},\, k \geqslant 1\}$ denote a family of random variables, defined on $(\Omega, \mathcal{F}, \mathbb{P})$ and with values in $\{0,1\}$. Then ΨH is the subset of $\mathbb{R} \times \mathbb{Z}$ defined as*

$$\Psi H := \{(\Psi(x,i,k), x)\,;\, x \in \mathbb{Z},\, i \in \mathfrak{I},\, k \geqslant 1,\, H(x,i,k) = 1\}.$$

Likewise,

$$\mathrm{Perf}(\xi, u, H) := \{\mathrm{Perf}(\xi, u, x, i, k) H(x,i,k)\,;\, x \in \mathbb{Z},\, i \in \mathfrak{I},\, k \geqslant 1\}\,.$$

Let $\mathcal{K}$ denote the product σ-algebra on $\{0,1\}^{\mathbb{Z}\times\mathfrak{I}\times\{1,2,\ldots\}}$.

Definition 2.10 (Coupling time with ambiguities)**.** *The pair (H,T) is a coupling time with ambiguities if $H = \{H(x,i,k)\,;\, x \in \mathbb{Z},\, i \in \mathfrak{I},\, k \geqslant 1\}$ is a $\{0,1\}$-valued process and T is a random variable defined on $(\Omega, \mathcal{F}, \mathbb{P})$, such that the following holds.*

1. *The random variable T belongs to $]-\infty, 0[$, $\mathbb{P}$ almost surely.*
2. *At most a finite number of the random variables $H(x,i,k)$ are not zero, $\mathbb{P}$ almost surely.*
3. *For every site x, rule index i and index $k \geqslant 1$, if $\Psi(x,i,k) < T$, then $H(x,i,k) = 0$, $\mathbb{P}$ almost surely.*
4. *For every site x, rule index i and index $k \geqslant 1$, $H(x,i,k)$ is measurable with respect to $\mathcal{F}^+(\Psi(x,i,k))$.*
5. *For every time $t < T$ and configurations ξ and ξ', if $\mathrm{Perf}(\xi,t,H)$ and $\mathrm{Perf}(\xi',t,H)$ are equal, then $\Phi(\xi,t,0^-,0)$ and $\Phi(\xi',t,0^-,0)$ are equal, $\mathbb{P}$ almost surely.*

Remark 2.11. *A consequence of definition 2.10 and of the flow property of Φ described by proposition 2.5 is that, if (H,T) is a coupling time with ambiguities, there exists a map $\Theta : \Omega \times \{0,1\}^{\mathbb{Z}\times\mathfrak{R}\times\{1,2,\ldots\}} \to S$ such that, $\mathbb{P}$ almost surely, for every $t < T$,*

$$\Phi(\xi, t, 0^-, 0) = \Theta(\Psi, \mathrm{Perf}(\xi, t, H)).$$

Note that Θ does not depend on t.

Definition 2.12 (Width of coupling times with ambiguities)**.** *The width of a coupling time with ambiguities (H,T) is bounded by the couple (a_-, a_+) of nonnegative integers if the following holds.*

1. *The random process $H = \{H(x,i,k)\,;\, x \in \mathbb{Z},\, i \in \mathfrak{I},\, k \geqslant 1\}$ is measurable with respect to $\sigma(\mathcal{F}(x,i)\,;\, -a_- \leqslant x \leqslant a_+,\, i \in \mathfrak{I})$.*
2. *$\mathrm{Influ}(\Psi H) \subset [-a_-, a_+] \times \mathbb{R}_-$, $\mathbb{P}$ almost surely.*
3. *The map Θ in remark 2.11 can be chosen to be measurable with respect to the σ-algebra $\sigma\left(\mathcal{F}(x,i)\,;\, -a_- \leqslant x \leqslant a_+,\, i \in \mathfrak{I}\right) \otimes \mathcal{K}$.*

Definition 2.13 (Growth parameter)**.** *The growth parameter of a coupling time with ambiguities (H,T) is*

$$m(H,T) := \mathbb{E}[\#\mathrm{Influ}(\Psi H)].$$

If $m(H,T) < 1$, we say that the coupling time with ambiguities (H,T) is subcritical.

Note that

$$\#\mathrm{Influ}(\Psi H) = \sum_{(x,i,k)} (\#A_i)\, H(x,i,k).$$

Definition 2.14 (Laplace transforms of coupling time with ambiguities)**.** *For every real number λ, introduce $\Lambda_T(\lambda) := \mathbb{E}[\mathrm{e}^{-\lambda T}]$ and*

$$\Lambda_H(\lambda) := \mathbb{E}\left(\sum_{(t,y)} \mathrm{e}^{-\lambda t}\mathbf{1}\{(t,y) \in \mathrm{Influ}(\Psi H)\}\right).$$

Recall that $T < 0$ almost surely and that $t < 0$ for every (t, y) in $\mathrm{Influ}(\Psi H)$. Note that $\Lambda_T(0) = 1$, $\Lambda_H(0) = m(H, T)$, and

$$\sum_{(t,y)} \mathrm{e}^{-\lambda t}\mathbf{1}\{(t,y) \in \mathrm{Influ}(\Psi H)\} = \sum_{(x,i,k)} (\#A_i)\, \mathrm{e}^{-\lambda\Psi(x,i,k)} H(x,i,k).$$

3. Statement of the main results

We are now able to state the main results of this paper.

Theorem A (Ergodicity)**.** *If there exists a subcritical coupling time with ambiguities, the particle system is ergodic. That is, there exists a unique invariant probability distribution μ, and, for every initial configuration ξ, X_t^ξ converges in distribution to μ as t goes to infinity.*

Theorems B and D below provide non-asymptotic results. Theorem C is a consequence of theorem B.

Theorem B (Explicit bound)**.** *Assume that there exists a subcritical coupling time with ambiguities (H, T) and let μ denote the unique invariant probability distribution of the particle system. For every configuration ξ, finite subset of sites $B \subset \mathbb{Z}$, and time $t \geqslant 0$, the distance in total variation between the distributions $\pi_B(X_t^\xi)$ and $\pi_B(\mu)$ is at most*

$$\#B \times \inf_{n\geqslant 0}\left(m(H,T)^n + \sum_{k=1}^{n} \inf_{\lambda\geqslant 0} \Lambda_H(\lambda)^k \Lambda_T(\lambda)\mathrm{e}^{-\lambda t}\right).$$

Theorem C (Exponential rate of convergence)**.** *Assume that there exists a subcritical coupling time with ambiguities (H, T) with finite width such that T is exponentially integrable. Then, with respect to the total variation distance, for every initial configuration ξ, the finite marginals of X_t^ξ converge exponentially fast to the finite marginals of the invariant distribution.*

Theorem D (Decay of correlations)**.** *Assume that there exists a subcritical coupling time with ambiguities with growth parameter $m < 1$ and finite width bounded by (a_-, a_+). Let μ denote the unique invariant probability distribution of the particle system. For every real number z, let $\kappa(z) := |z|/(a_+ + a_-)$.*
For every sites x and y in $\mathbb{Z}$, the distance in total variation between $\pi_{\{x,y\}}(\mu)$ and $\pi_x(\mu) \otimes \pi_y(\mu) = \pi_0(\mu) \otimes \pi_0(\mu)$ is at most $2m^{\kappa(y-x)-1}$.

Let x denote a positive integer, $B := \mathbb{Z} \cap [x, +\infty)$ and $C := \mathbb{Z} \cap (-\infty, 0]$. The distance in total variation between $\pi_{B\cup C}(\mu)$ and $\pi_B(\mu) \otimes \pi_C(\mu)$ is at most

$$m^{\kappa(x)-1}\left(\frac{1}{1-m^{1/a_-}} + \frac{1}{1-m^{1/a_+}}\right).$$

As a consequence, the same bound applies to every subsets B and C of $\mathbb{Z}$ such that $\min B \geqslant x + \max C$.

4. Proof of the main results

In section 4.1, we define a crucial tool for our proofs, namely the notion of ambiguity processes. In section 4.2, we explain how to control these. This enables us to prove theorem A in section 4.3, theorem B in section 4.4, and theorem D in section 4.5. Section 4.6 is a preparation to the proof of theorem C, given in section 4.7. Finally, section 4.8 settles some measurability issues.

4.1. **Ambiguity processes.**

For every sites x and y, time $t \leqslant 0$, rule index i in $\mathfrak{I}$ and index $k \geqslant 1$, let

$$H(t,x)(y,i,k) := (H \circ \sigma_{t,x})(y-x, i, k - K_t(y,i)).$$

For every site x and time $t \leqslant 0$, let

$$H(x,t) := \{H(x,t)(y,i,k)\,;\, y \in \mathbb{Z},\, i \in \mathfrak{I},\, k \geqslant 1\}.$$

Similarly, let

$$T(x,t) := (T \circ \sigma_{t,x}) + t.$$

In words, the pair $(H(t,x), T(t,x))$ corresponds to the translation of the coupling time with ambiguities (H,T) from site 0 and time 0^- to site x and time t^-.

Definition 4.1. *For every site x and time $t \leqslant 0$, let $A(t,x)$ denote the event that, for every time $u < T(x,t)$ and configurations ξ and ξ' such that $\mathrm{Perf}(\xi, u, H(t,x))$ and $\mathrm{Perf}(\xi', u, H(t,x))$ coincide, $\Phi(\xi, u, t^-, x)$ and $\Phi(\xi', u, t^-, x)$ coincide.*

Let Ω_3 be the event

$$\Omega_3 := A(0,0) \cap \bigcap_{(x,y,i,k)} A(\Psi(y,i,k), x).$$

Lemma 4.2. ***(1)*** *For every site x and time $t \leqslant 0$, $\mathbb{P}[A(t,x)] = 1$.*
(2) *For every sites x and y, rule index i and index $k \geqslant 1$,*

$$\mathbb{P}[A(\Psi(y,i,k), x)] = 1.$$

Proof of lemma 4.2. Part (1) is a simple consequence of lemma 2.7 in section 2.3.3, and of the fact that $\mathbb{P}$ is invariant under the action of $\sigma_{t,x}$. We omit the details of the proof.

As regards part (2), by our definition 2.10 above, there exists a set B in $\mathcal{F}$ such that $B \subset A(0,0)$ and $\mathbb{P}[B] = 1$. With our definitions, for every x and t,

$$\Omega_1 \cap \Omega_2 \cap \{\psi \in \Omega\,;\, \sigma_{t,x}(\psi) \in B\} \subset A(t,x).$$

Hence, we only need to prove that $\mathbb{P}[\sigma_{\tau,x}(\Psi) \in B] = 1$, where $\tau := \Psi(y,i,k)$.

By standard properties of Poisson processes, $\sigma_{\tau,x}(\Psi)$ is independent from τ and has the same distribution as Ψ. As a consequence, letting $\mathbb{T}$ denote the distribution of τ,

the distribution of $(\tau, \sigma_{\tau,x}(\Psi))$ on $(-\infty, 0) \times \Omega$ equipped with the product σ-algebra $\mathcal{B}((-\infty, 0)) \otimes \mathcal{F}$, is equal to the product measure $\mathbb{T} \otimes \mathbb{P}$. By Fubini theorem,

$$\mathbb{P}[\sigma_{x,\tau}(\Psi) \in B] = \int_{-\infty}^{0} \mathbb{P}[\Psi \in B] \, \mathrm{d}\mathbb{T}(t) = 1.$$

This concludes the proof of lemma 4.2. □

A direct consequence of lemma 4.2 above is the following proposition.

Proposition 4.3. $\mathbb{P}[\Omega_3] = 1$.

Definition 4.4 (Ambiguity processes)**.** *The ambiguity process at site x is a sequence $(\mathrm{Amb}_n(x))_{n \geqslant 0}$ of random subsets of $\mathbb{Z} \times \mathbb{R}_-$, defined recursively as follows.*

- *Initialization:* $\mathrm{Amb}_0(x) := \{(0, x)\}$.
- *Induction:* $\mathrm{Amb}_{n+1}(x) := \bigcup_{(t,y) \in \mathrm{Amb}_n(x)} \mathrm{Influ}(\Psi H(t, y))$.

This defines a nondecreasing random sequence of sets $(\mathrm{Amb}_n(x))_{n \geqslant 0}$. Let

$$\mathrm{Amb}(x) := \bigcup_{n \geqslant 0} \mathrm{Amb}_n(x).$$

By construction, $\mathrm{Amb}(x)$ is a finite set if and only $\mathrm{Amb}_n(x) = \mathrm{Amb}_{n+1}(x)$ for some index $n \geqslant 0$. For such an index n, $\mathrm{Amb}_k(x) = \mathrm{Amb}_n(x)$ for every $k \geqslant n$, whence $\mathrm{Amb}_n(x) = \mathrm{Amb}(x)$.

We wish to prove that for subcritical coupling times with ambiguities, the set $\mathrm{Amb}(x)$ is almost surely finite.

Definition 4.5 (Coupling time at a site)**.** *The coupling time T_x^* at site x is*

$$T_x^* := \inf\{T(t, y) \, ; \, (t, y) \in \mathrm{Amb}(x)\}.$$

Observe that, if $\mathrm{Amb}_n(x) = \mathrm{Amb}_{n+1}(x)$, then

$$T_x^* := \inf\{T(t, y) \, ; \, (t, y) \in \mathrm{Amb}_n(x)\}.$$

Definition 4.6 (Influence in the ambiguity process)**.** *Let (t, y) and (u, z) denote elements of $\mathrm{Amb}(x)$. Say that (u, z) is influenced by (t, y) if (u, z) belongs to the set $\mathrm{Influ}(\Psi H(t, y))$.*

Note that an element of $\mathrm{Amb}(x)$ may be influenced by several elements of $\mathrm{Amb}(x)$.

We define inductively a sequence $(Z_n(x))_{n \geqslant 0}$ of random subsets of $\mathrm{Amb}(x)$ as follows. For $n = 0$, let $Z_0(x) := \{(0, x)\}$. For every $n \geqslant 0$, $Z_{n+1}(x)$ denotes the set (possibly empty) of the elements influenced by elements of $Z_n(x)$. Hence,

$$\mathrm{Amb}(x) = \bigcup_{n \geqslant 0} Z_n(x).$$

Moreover, if $Z_n(x)$ is empty for a given $n \geqslant 1$, then $Z_k(x)$ is empty for every $k \geqslant n$ as well, and in that case,

$$\mathrm{Amb}(x) = \bigcup_{0 \leqslant k \leqslant n-1} Z_k(x).$$

Definition 4.7 (Locked sites)**.** *For every site x and times $u < t \leqslant 0$, say that (t^-, x) is locked by time u if for every time $v < u$ and configurations ξ and ξ',*

$$\Phi(\xi, v, t^-, x) = \Phi(\xi', v, t^-, x).$$

Lemma 4.8. *On $\Omega_1 \cap \Omega_2$, if (t^-, x) is locked by time $u < t$, then for every times $v < u$ and $v' < u$, and configurations ξ and ξ',*

$$\Phi(\xi, v, t^-, x) = \Phi(\xi', v', t^-, x)$$

The proof of lemma 4.8 is a consequence of the flow property of Φ in proposition 2.5 and we omit it.

We now give a definition concerning ambiguities.

Definition 4.9 (Resolution of ambiguities)**.** *One says that the ambiguity associated to $(\Psi(x, i, k), x)$ is resolved by time u if $u < \Psi(x, i, k)$ and if, for every configurations ξ and ξ',*

$$\mathrm{Perf}(\xi, u, x, i, k) = \mathrm{Perf}(\xi', u, x, i, k).$$

Proposition 4.10. *If $\mathrm{Amb}(x)$ is finite, then $(0^-, x)$ is almost surely locked by time t, for every $t < T_x^*$.*

Proof of proposition 4.10. Assume throughout the proof that $\Omega_1 \cap \Omega_2 \cap \Omega_3$ holds, since this event has probability one.

Consider an index n such that $\mathrm{Amb}_x(n) = \mathrm{Amb}_x(n+1)$, and let $(t_k, x_k)_{1 \leqslant k \leqslant r}$ denote an enumeration of the set $(\mathrm{Amb}_x(n)) \cap]-\infty, 0]$ such that

$$t_1 < t_2 < \cdots < t_r = 0.$$

We wish to prove by induction that, for every $1 \leqslant q \leqslant r$, with full probability, the following property (P_q) holds:

> (P_q) For every $1 \leqslant j \leqslant q$, and every (t_j, y) in $\mathrm{Influ}(t_j, x_j)$, (t_j^-, y) is locked by time T_x^*.

Assume first that $q = 1$ and consider (t_1, z) in $\mathrm{Influ}(t_1)$.

By definition, $\mathrm{Influ}(\Psi H(t_1, z)) \subset \mathrm{Amb}_x(n+1)$ since (t_1, z) belongs to $\mathrm{Amb}_x(n)$. But every element (w, u) in $\mathrm{Influ}(\Psi H(t_1, z))$ is such that $u < t_1$, by definition of H and Influ. On the other hand, $u > t_1$ by the definition of t_1. This is a contradiction, hence $\mathrm{Influ}(\Psi H(t_1, z))$ is empty. Using the definition of a coupling time with ambiguities and the fact that, by definition, $T_x^* \leqslant T(t_1, z)$, we deduce that (t_1^-, z) is, with full probability, locked by time T_x^*. This proves (P_1).

Assume now that (P_q) hold for some $1 \leqslant q \leqslant r-1$, and consider an element (t_{q+1}, z) of $\mathrm{Influ}(t_{q+1}, x_{q+1})$. Observe that $H(t_{q+1}, z) \subset \{(t_1, x_1), \ldots, (t_q, x_q)\}$. As a consequence, according to (P_q), with full probability, the ambiguities associated with the elements of $H(t_{q+1}, z)$ are resolved by time T_x^*. Thus, for every configurations ξ and ξ' and time $t < T_x^*$,

$$\mathrm{Perf}(t, \xi, H(t_{q+1}, z)) = \mathrm{Perf}(t, \xi', H(t_{q+1}, z)).$$

Using the fact that by definition $T_x^* \leqslant T(t_{q+1}, z)$, and the definition of a coupling time with ambiguities, one sees that (t_{q+1}^-, z) is locked by time T_x^*. Hence the ambiguity associated with (t_{q+1}, x_{q+1}) is resolved by time T_x^*, and (P_{q+1}) holds.

The proof of proposition 4.10 is complete. □

Remark 4.11. *The reader might have noticed that property (4) in definition 2.10 in section 2.4 is not used in the proof of proposition 4.10 above. However, this property plays a crucial role in the estimates on the ambiguity process presented in the next section.*

4.2. **Controlling ambiguity processes.** The goal of this section is to prove the preliminary estimates of lemma 4.12 below. For every λ and nonnegative integer n, let

$$C(n,\lambda) := \sum_{(t,y)\in Z_n(0)} \mathrm{e}^{-\lambda t},$$

and

$$D(n,\lambda) := \sum_{(t,y)\in Z_n(0)} \mathrm{e}^{-\lambda T(t,y)}.$$

Lemma 4.12. *For every λ and nonnegative integer n,*

$$\mathbb{E}[C(n,\lambda)] \leqslant \Lambda_H(\lambda)^n,$$

and

$$\mathbb{E}[D(n,\lambda)] = \Lambda_T(\lambda)\mathbb{E}(C(n,\lambda)) \leqslant \Lambda_T(\lambda)\Lambda_H(\lambda)^n.$$

Proof of lemma 4.12. The proof of the first assertion is by induction on n. For $n = 0$, $C(0,\lambda) = 1$ hence the result is obvious. Assume that the result holds for a given $n \geqslant 0$. Every element in $Z_{n+1}(x)$ is influenced by at least one element in $Z_n(x)$, hence $C(n+1,\lambda) \leqslant C_0(n+1,\lambda)$, with

$$C_0(n+1,\lambda) := \sum_{(t,y)\in Z_n(0)} \sum_{(u,z)} \mathrm{e}^{-\lambda u} \mathbf{1}\{(u,z) \in \mathrm{Influ}(\Psi H(t,y))\}.$$

Hence,

$$C_0(n+1,\lambda) = \sum_{(t,y)\in Z_n(0)} \mathrm{e}^{-\lambda t} \sum_{(u,z)} \mathrm{e}^{-\lambda(u-t)} \mathbf{1}\{(u,z) \in \mathrm{Influ}(\Psi H(t,y))\}.$$

Let

$$C_1(n,\lambda,x,y,i,k) := \mathbf{1}\{(y,\Psi(x,i,k)) \in Z_n(0)\}\, \mathrm{e}^{-\lambda\Psi(x,i,k)},$$

and

$$C_2(\lambda,x,y,i,k) := \sum_{(u,z)} \mathrm{e}^{-\lambda(u-\Psi(x,i,k))}\, \mathbf{1}\{(u,z) \in \mathrm{Influ}(\Psi H(\Psi(x,i,k),y))\}.$$

The last expression of $C_0(n+1,\lambda)$ can be rewritten as

$$C_0(n+1,\lambda) = \sum_{(x,i,k)} \sum_{y\in x+A_i} C_1(n,\lambda,x,y,i,k) C_2(\lambda,x,y,i,k).$$

Taking expectations on both sides,

$$\mathbb{E}[C_0(n+1,\lambda)] = \sum_{(x,i,k)} \sum_{y\in x+A_i} \mathbb{E}\left[C_1(n,\lambda,x,y,i,k) C_2(\lambda,x,y,i,k)\right].$$

According to lemmas 4.16 and 4.17 in section 4.8 below, every $C_1(n,\lambda,x,y,i,k)$ is measurable with respect to $\mathcal{F}^+(\Psi(x,i,k))$, while the conditional distribution of every $C_2(\lambda,x,y,i,k)$ with respect to $\mathcal{F}^+(\Psi(x,i,k))$ is the same as the (unconditional) distribution of

$$\sum_{(u,z)} \mathrm{e}^{-\lambda u} \mathbf{1}\{(u,z) \in \mathrm{Influ}(\Psi H)\}.$$

As a consequence,

$$\mathbb{E}\left[C_1(n,\lambda,x,y,i,k)C_2(\lambda,x,y,i,k)\right] = \Lambda_H(\lambda)\,\mathbb{E}[C_1(n,\lambda,x,y,i,k)].$$

In turn, this implies that

$$\mathbb{E}[C_0(n+1,\lambda)] = \Lambda_H(\lambda)\mathbb{E}\left(\sum_{(x,i,k)}\sum_{y\in x+A_i} C_1(n,\lambda,x,y,i,k)\right).$$

It remains to notice that

$$\sum_{(x,i,k)}\sum_{y\in x+A_i} C_1(n,\lambda,x,y,i,k) = C(n,\lambda),$$

to see that the induction on n is complete. The proof of the first assertion of lemma 4.12 is complete.

As regards the second assertion, fix an integer $n \geqslant 0$, and define

$$C_3(\lambda,x,y,i,k) := \mathrm{e}^{-\lambda(T(\Psi(x,i,k))-\Psi(x,i,k),y)}.$$

Using the functional C_1 defined in the proof of the first assertion, one sees that

$$D(n,\lambda) = \sum_{(x,i,k)}\sum_{y\in x+A_i} C_1(n,\lambda,x,y,i,k)C_3(\lambda,x,y,i,k).$$

As a consequence,

$$\mathbb{E}[D(n,\lambda)] = \sum_{(x,i,k)}\sum_{y\in x+A_i} \mathbb{E}[C_1(n,\lambda,x,y,i,k)C_3(\lambda,x,y,i,k)].$$

As in the proof of the first assertion, we observe that, for every fixed (x,i,k,y), the random variable $C_1(n,\lambda,x,y,i,k)$ is measurable with respect to $\mathcal{F}^+(\Psi(x,i,k))$, while the conditional distribution of $T(y,\Psi(x,i,k)) - \Psi(x,i,k)$ with respect to $\mathcal{F}^+(\Psi(x,i,k))$, is the same as the (unconditional) distribution of T. Using the last displayed identity in the proof of the first assertion once again, the result follows.

The proof of lemma 4.12 is complete. □

4.3. **End of the proof of theorem A.** Lemma 4.12 with $\lambda = 0$ shows that, for every $n \geqslant 0$ and every site x, $\mathbb{E}[\#Z_n(x)] \leqslant \Lambda_H(0)^n = m(H,T)^n$. Markov inequality yields $\mathbb{P}[Z_n(x) \neq \emptyset] \leqslant m(H,T)^n$. In particular, there exists $\mathbb{P}$ almost surely an integer n such that $Z_n(x)$ is empty. For such an integer n,

$$\mathrm{Amb}(x) = \bigcup_{0\leqslant k\leqslant n-1} Z_k(x),$$

hence $\mathrm{Amb}(x)$ is finite with full probability.

Fix a finite subset $B \subset \mathbb{Z}$, and let $T_B^* := \min\{T_x^*\,;\, x \in B\}$. By proposition 4.10 in section 4.1, for every time $t < T_x^*$ and configurations ξ and ξ',

$$\pi_B(\Phi(\xi,t,0^-,x)) = \pi_B(\Phi(\xi',t,0^-,x)).$$

The probability that time 0 belongs to Ψ is 0, hence, almost surely, for every time $t < T_x^*$ and configurations ξ and ξ',

$$\pi_B(\Phi(\xi,t,0,x)) = \pi_B(\Phi(\xi',t,0,x)).$$

As a consequence, for every positive time t and configurations ξ and ξ', the distance in total variation between $\pi_B(X_t^\xi)$ and $\pi_B(X_t^{\xi'})$ is at most $\mathbb{P}[T_B^* < -t]$. Since every

T_x^* is almost surely finite and B is finite, T_B^* is almost surely finite, hence this distance goes to 0 when t goes to infinity.

Now, a generic compactness argument shows that, as a consequence of S being finite, there exists at least one invariant distribution μ for the particle system (see for instance [10] chapter 1, proposition 1.8). Consider now a random configuration ξ' with distribution μ, and an arbitrary configuration ξ. For every $t \geqslant 0$, the distribution of $X_t^{\xi'}$ is μ, and the previous estimates then show that, as t goes to infinity, X_t^{ξ} converges in distribution to μ as t goes to infinity. This ends the proof of theorem A.

4.4. **End of the proof of theorem B.** By the union bound, for every negative t,

$$\mathbb{P}[T_B^* \leqslant t] \leqslant \sum_{x \in B} \mathbb{P}[T_x^* \leqslant t] = (\#B)\, \mathbb{P}[T_0^* \leqslant t].$$

Fix an index $n \geqslant 1$. On the event that $Z_n(0)$ is empty,

$$T_0^* = \min\{T(u,x)\,;\, 0 \leqslant k \leqslant n-1,\, (u,x) \in Z_k(0)\}.$$

As a consequence, by the union bound,

$$\mathbb{P}[T_0^* \leqslant t] \leqslant m(H,T)^n + \sum_{k=0}^{n-1} \mathbb{P}\left[\min_{(u,x)\in Z_k(0)} T(u,x) \leqslant t\right].$$

Fix $\lambda \geqslant 0$. Then

$$\exp(-\lambda \min_{(u,x)\in Z_k(0)} T(u,x)) \leqslant D(k,\lambda).$$

Hence, Markov inequality yields

$$\mathbb{P}\left[\min_{(u,x)\in Z_k(0)} T(u,x) \leqslant t\right] \leqslant \mathrm{e}^{\lambda t}\mathbb{E}[D(k,\lambda)] \leqslant \mathrm{e}^{\lambda t}\Lambda_H(\lambda)^k \Lambda_T(\lambda),$$

where lemma 4.12 provides the last inequality. Taking the infimum with respect to $\lambda \geqslant 0$ in this inequality, separately for each $0 \leqslant k \leqslant n-1$, yields the conclusion of theorem B.

4.5. **Proof of theorem C.** By hypothesis, $\Lambda_T(\lambda)$ is finite for some positive values of λ and the width of (H,T) is finite, hence our next lemma implies the theorem.

Lemma 4.13. *Assume that the width of (H,T) is finite and that $\Lambda_T(\lambda)$ is finite for a positive λ. Then $\Lambda_H(\lambda')$ is finite for every $\lambda' < \lambda$.*

Proof of lemma 4.13. Replacing every $H(x,i,k)$ such that $-a_- \leqslant x \leqslant a_+$ and $\Psi(x,i,k) \geqslant T$ by 1 in the expectation which defines $\Lambda_H(\lambda')$ yields

$$\Lambda_H(\lambda') \leqslant \sum_{-a_- \leqslant x \leqslant a_+} \sum_i \mathbb{E}[\Lambda_H(x,i,\lambda')],$$

with

$$\Lambda_H(x,i,\lambda') := \sum_{t \in \Psi(x,i)} \mathrm{e}^{-\lambda' t}\mathbf{1}\{t > T\}.$$

Fix a positive real number λ''. Then, $\mathrm{e}^{-\lambda' t}\mathbf{1}\{t > T\} \leqslant \mathrm{e}^{-\lambda'' T + (\lambda''-\lambda')t}$, hence, for every site x and rule index i,

$$\Lambda_H(x,i,\lambda') \leqslant \mathrm{e}^{-\lambda'' T} \sum_{t \in \Psi(x,i)} \mathrm{e}^{(\lambda''-\lambda')t}.$$

Fix $p > 1$ and $q > 1$ such that $1/p + 1/q = 1$. By Minkowski inequality,

$$\mathbb{E}[\Lambda_H(x,i,\lambda')] \leqslant \mathbb{E}[\mathrm{e}^{-p\lambda'' T}]^{1/p} \sum_{k \geqslant 1} \mathbb{E}[\mathrm{e}^{q(\lambda''-\lambda')\Psi(x,i,k)}]^{1/q}.$$

Since each $-\Psi(x,i,k)$ is the sum of k i.i.d. exponential random variables, the last sum is the sum of a geometric series, with ratio $\mathbb{E}[\mathrm{e}^{q(\lambda''-\lambda')\Psi(x,i,1)}]^{1/q}$. Assume that $\lambda'' > \lambda'$. The ratio is less than 1, hence the sum over $k \geqslant 1$ converges.

Summing this over every site $-a_- \leqslant x \leqslant a_+$ and rule index i yields a finite upper bound of $\Lambda_H(\lambda')$ as soon as one can find $\lambda'' > \lambda'$ and (p,q) such that $1/p + 1/q = 1$ and $\mathbb{E}[\mathrm{e}^{-p\lambda'' T}]$ is finite. If $\Lambda_T(\lambda) = \mathbb{E}[\mathrm{e}^{-\lambda T}]$ is finite, this is possible for every $\lambda' < \lambda$, hence the proof of lemma 4.13 is complete. □

4.6. **Preparation to the proof of theorem D.**

Definition 4.14. *For every site x, let $N^*(x)$ denote the, almost surely finite, smallest integer n such that $Z_n(x)$ is empty, and let $\mathcal{G}_x$ denote the σ-algebra*

$$\mathcal{G}_x := \sigma(\mathcal{F}(y,i)\,;\, x + a_- N^*(x) \leqslant y \leqslant x + a_+ N^*(x),\, i \in \mathfrak{I}).$$

Lemma 4.15. *For every site x, there exists a random variable F_x, with values in S, which is measurable with respect to $\mathcal{G}_x$ and such that, $\mathbb{P}$ almost surely and for every configuration ξ,*

$$\lim_{t \to -\infty} \Phi(\xi, t, 0, x) = F_x.$$

Proof of lemma 4.15. Assume throughout the proof that $\Omega_1 \cap \Omega_2 \cap \Omega_3$ holds, since this event has probability one. Let $(t_k, x_k)_{1 \leqslant k \leqslant r}$ denote an enumeration of the set $\mathrm{Influ}^{-1}(\mathrm{Amb}_x(N^*(x))) \cap]-\infty, 0]$ such that $t_1 < t_2 < \cdots < t_r = 0$.

For every $1 \leqslant m \leqslant r-1$, define (x_m, i_m, k_m) by the relation

$$t_m =: \Psi(x_m, i_m, k_m).$$

We define by induction a sequence $(W_m)_{0 \leqslant i \leqslant r-1}$ such that, for every m,

$$W_m \in \{0,1\}^{\mathbb{Z} \times \mathfrak{I} \times \{1,2,\ldots\}}.$$

First, let $W_0(x,i,k) := 0$ for every x, i and k. Now, let $1 \leqslant m \leqslant r-1$. For every $q \leqslant m-1$, let $W_m(x_q, i_q, k_q) := W_{m-1}(x_q, i_q, k_q)$. For $q = m$, let

$$W_m(x_m, i_m, k_m) := \mathbf{1}\left\{[\Theta(\sigma_{t_m, y}(\Psi)), W_{m-1} H(t_m, x_m)\,;\, y \in x_m + A_m] \in \ell_m\right\}.$$

Finally, let $W_m(x,i,k) := 0$ when (x,i,k) is not one of the triples (x_j, i_j, k_j) for $1 \leqslant j \leqslant m$.

As in the proof of proposition 4.10, we have that, for every $q \leqslant r-1$, $\mathbb{P}$ almost surely, for every time $t < T_x^*$ and for every (t_{q+1}, y) in $\mathrm{Influ}(t_{q+1}, x_{q+1})$, $\Psi H(t_{q+1}, y)$ is a subset of $\{(t_1, x_1), \ldots, (t_q, x_q)\}$, and that

$$\Phi(\xi, t, t_{q+1}^-, y) = \Theta(\mathrm{Perf}(\xi, t, H(t_{q+1}, y)), \sigma_{t_{q+1}, y}(\Psi)).$$

Hence $\mathrm{Perf}(\xi, t, H(t_q, y)) = W_{q-1} H(t_q, y)$ for every $1 \leqslant q \leqslant r$ and $\mathbb{P}$ almost surely. Let

$$F_x := \Theta(W_{r-1} H(x, 0), \sigma_{0,x}(\Psi)).$$

One sees that, $\mathbb{P}$ almost surely and for every time $t < T_x^*$,

$$\Phi(\xi, t, 0, x) = \Phi(\xi, t, 0^-, x) = F_x.$$

Finally, the measurability properties of F_x follow from assumption (3) in definition 2.12. The proof of lemma 4.15 is complete. □

4.7. **Proof of theorem D.** Assume that $\min B = n + \max C$ with $n \geqslant 1$, and introduce the real number

$$z := \max C + n\frac{a_+}{a_+ + a_-} = \max C + \kappa(n)a_+.$$

Let Ψ_C and Ψ_B denote two families of Poisson processes indexed by $\mathbb{Z}$ and such that the following properties hold.

- $\Psi_C(x) = \Psi(x)$ for every $x < z$ and $\Psi_B(x) = \Psi(x)$ for every $x > z$.
- $(\Psi_C(x))_{x\geqslant z}$ has the same distribution as $(\Psi(x))_{x\geqslant z}$ but is independent from (Ψ, Ψ_B).
- $(\Psi_B(x))_{x\leqslant z}$ has the same distribution as $(\Psi(x))_{x\leqslant z}$ but is independent from (Ψ, Ψ_C).

Hence Ψ_B and Ψ_C are independent. Recall lemma 4.15 and definition 4.14 in section 4.6 and let

$$N_C = \sup\{x + a_+N^*(x);\ x \in C\},$$

and

$$N_B = \inf\{x - a_-N^*(x);\ x \in B\}.$$

Consider the event

$$E := \{N_C < z < N_B\}.$$

On E, $((F_x(\Psi))_{x\in C}, (F_x(\Psi))_{x\in B})$ is distributed like $((F_x(\Psi_C))_{x\in C}, (F_x(\Psi_B))_{x\in B})$. Since Ψ_C and Ψ_B are independent, $(F_x(\Psi))_{x\in C}$ and $(F_x(\Psi))_{x\in B}$ are independent on E. Hence the distance in total variation we want to estimate is at most $\mathbb{P}[\Omega \setminus E]$.

Note that $\Omega \setminus E = \{N_B \leqslant z\} \cup \{N_C \geqslant z\}$. As regards N_B,

$$\{N_B \leqslant z\} = \bigcup_{x\in B} \{x - a_-N^*(x) \leqslant z\},$$

hence

$$\mathbb{P}[N_B \leqslant z] \leqslant \sum_{x\geqslant \min B} \mathbb{P}[x - a_-N^*(x) \leqslant z].$$

For each $x \geqslant \min B$, $\{x - a_-N^*(x) \leqslant z\} = \{N^*(x) \geqslant \kappa(n) + k/a_-\}$, provided k is the nonnegative integer $k := x - \min B$. Furthermore, for every real number v, $\mathbb{P}[N^*(x) \geqslant v] \leqslant m^{v-1}$ with $m := m(H, T) < 1$. This yields

$$\mathbb{P}[N_B \leqslant z] \leqslant \sum_{k\geqslant 0} m^{\kappa(n)-1+k/a_-} = \frac{m^{\kappa(n)-1}}{1 - m^{1/a_-}}.$$

The same argument, applied to $\mathbb{P}[N_C \geqslant z]$, and the union bound, yield the result.

The case when $C = \{x\}$ and $B = \{y\}$ with $y > x$ is similar, except that one can replace the geometric series involved above by their first term, hence the tighter bounds.

The proof of theorem D is complete.

4.8. **Measurability properties.** Let $\mu(H)$ denote the random counting measure associated to ΨH, that is,

$$\mu(H) := \sum_{(x,i,k)} \mathbf{1}\{H(x,i,k) = 1\}\, \delta_{\Psi(x,i,k)}.$$

Lemma 4.16. *Fix sites x and y, a rule index i, an index $k \geqslant 1$, and a nonnegative measurable function $F : \mathbb{R} \to \mathbb{R}_+$. Then the random variable $\displaystyle\int F \mathrm{d}\mu(H \circ \sigma_{\Psi(x,i,k),y})$ is independent from $\mathcal{F}^+(\Psi(x,i,k))$ and has the same distribution as $\displaystyle\int F \mathrm{d}\mu(H)$.*

Proof of lemma 4.16. The independence property is a consequence of the independence of $\sigma_{\Psi(x,i,k),y}$ and $\mathcal{F}^+(\Psi(x,i,k))$. The equidistribution property is a consequence of the invariance of $\mathbb{P}$ with respect to $\sigma_{\Psi(x,i,k),y}$. □

Lemma 4.17. *For every integer $n \geqslant 0$, the event $\{\Psi(x,i,k) \in Z_n(x)\}$ is measurable with respect to $\mathcal{F}^+(\Psi(x,i,k))$.*

Proof of lemma 4.17. We proceed by induction on $n \geqslant 0$. The case of $n = 0$ is included in the definition of a coupling time with ambiguities. Let $n \geqslant 1$. By definition, $\{\Psi(x,i,k) \in Z_n(x)\}$ is the union of the events

$$E(x',i',k',y) := \{\Psi(x',i',k') \in Z_{n-1}(x),\ \Psi(x,i,k) \in H(y, \Psi(x',i',k'))\},$$

over every site x', rule index i', index k' and site y in $x' + A_{i'}$.

In the rest of this proof, we use τ and τ' as shorthands for $\tau := \Psi(x,i,k)$ and $\tau' := \Psi(x',i',k')$, respectively.

Lemma 4.17 follows from claims 4.18 and 4.19 below.

Claim 4.18. *For every event G in $\mathcal{F}^+(\tau')$, $G \cap \{\tau < \tau'\}$ belongs to $\mathcal{F}^+(\tau)$.*

Claim 4.19. *The event $\{\tau \in H(\tau', y)\}$ belongs to $\mathcal{F}^+(\tau)$.*

Indeed, by the induction hypothesis, $\{\tau' \in Z_{n-1}(x)\}$ belongs to $\mathcal{F}^+(\tau')$. Hence, using claim 4.18, we see that $\{\tau' \in Z_{n-1}(x)\} \cap \{\tau < \tau'\}$ belongs to $\mathcal{F}^+(\tau)$. On the other hand, by claim 4.19, we see that $\{\tau \in H(\tau', y)\}$ belongs to $\mathcal{F}^+(\tau)$ too. Since, by definition, $\{\tau \in H(\tau', y)\} \subset \{\tau < \tau'\}$, we finally deduce that any event $E(x', i', k', y)$ defined above belongs to $\mathcal{F}^+(\tau)$. This ends the proof of lemma 4.17. □

Proof of claim 4.18. It is easily checked, using the fact that $\{\tau < \tau'\}$ belongs to $\mathcal{F}^+(\tau)$, that the family of the events G which share the property stated in the claim is a σ-algebra. By definition, $\mathcal{F}^+(\tau')$ is generated by events of the form

$$\{K_{\tau'}(x_1, i_1) = k_1\} \cap \{\Psi(x_1, i_1, k_2) \in B\},$$

for every site x_1, rule index i_1, and indices k_1 and k_2 such that $k_2 \leqslant k_1$, and every Borel subset B of $\mathbb{R}$.

Now for every event G of this form, it is easy to check that $G \cap \{\tau < \tau'\}$ belongs to $\mathcal{F}^+(\tau)$. Claim 4.18 follows. □

Proof of claim 4.19. Observe that $\{\tau \in H_{y,\tau'}\}$ is the union over every $1 \leqslant m \leqslant k-1$ of the events

$$\{H(x-y,i,k-m) \circ \sigma_{\tau',y} = 1\} \cap \{K_{\tau'}(x,i) = m\} \cap \{\tau < \tau'\}.$$

Fix $m \leqslant k-1$. The collection of sets G in $\mathcal{F}^+(\Psi(x-y,i,k-m))$ such that the event

$$\{\mathbf{1}_G \circ \sigma_{\tau',y} = 1\} \cap \{K_{\tau'}(x,i) = m\} \cap \{\tau < \tau'\}$$

belongs to $\mathcal{F}^+(\tau)$ is a σ-algebra. On the other hand, the σ-algebra $\mathcal{F}^+(\tau)$ is generated by the events of the form

$$\{K_\tau(x_1,i_1) = k_1\} \cap \{\Psi(x_1,i_1,k_2) \in B\},$$

for every site x_1, rule index i_1, and integers k_1 and k_2 such that $k_2 \leqslant k_1$, and every Borel subset B of $\mathbb{R}$. For every event G of this form, the definition of a coupling time with ambiguities yields the fact that the event

$$\{\mathbf{1}_G \circ \sigma_{\tau',y} = 1\} \cap \{K_{\tau'}(x,i) = m\} \cap \{\tau < \tau'\}$$

belongs to $\mathcal{F}^+(\tau)$. Claim 4.19 follows. □

5. Applications to nucleotide substitution models

Most stochastic models of nucleotidic substitution processes assume that the various sites along a DNA sequence evolve independently. However, it is a well-known experimental fact that the nucleotides in the immediate neighborhood of a site can affect drastically the substitution rates at this site. For instance, in the genomes of vertebrates, the increased rates of substitution of cytosine by thymine and of guanine by adenine in CpG dinucleotides are often quite noticeable (typical ratios 10:1 when compared to the other rates of substitution). Recently, various models that take such dependences into account have been proposed, see [2, 4, 5, 7, 8, 12, 17] for instance. Among these, the class of RN+YpR models of nucleotide substitution, introduced by molecular biologists, and studied mathematically in [3], enjoys some remarkable properties, such as the possibility to solve exactly for several quantities of interest, and the occurrence of a non-zero but finite-range dependence along the sequence.

Since these models put restrictive conditions on substitution rates (see below) that may be only approximately satisfied in some actual biological situations of interest, it is biologically relevant to study the properties of nucleotide substitution models that are not in the RN+YpR class but close to some models in this class. From a mathematical perspective, it is interesting to study what becomes of the dependencies along the sequence when small perturbations of the RN+YpR assumptions are introduced, thus destroying the special mechanism leading to finite-range dependence in the RN+YpR context.

We apply the coupling techniques described in the rest of the paper to a generic family of perturbations of models in this class. In section 5.1, we describe the class of models. In section 5.2, we introduce two coupling times with ambiguities and we state theorem E, our main result about these. The proof of theorem E is in section 5.3. In section 5.4, we state some remarks. In section 5.5, we compute the growth parameters associated to these coupling times with ambiguities. The computations are based on some tree decompositions of conditional distributions,

stated in section 5.6. Finally, in section 5.7, we apply these results to the simplest non trivial example, namely the perturbed Jukes-Cantor model with CpG influence, thus proving the quantitative result stated as theorem F.

We mention that the notations in this section are sometimes slightly at odds with those in the rest of the paper.

5.1. **Description of the models.**

5.1.1. *RN+YpR models.* Formally, these models are interacting particle systems with state space $\mathcal{A}^{\mathbb{Z}}$, where

$$\mathcal{A} := \{A, T, C, G\}$$

denotes the nucleotidic alphabet. The letters A and G correspond to purines, abbreviated collectively by R, while C and T correspond to pyrimidines, abbreviated collectively by Y. Such a model is characterized by two sets of parameters, which describe two distinct kinds of transition mechanisms.

The RN part of the model consists of a matrix $S := (s_{x,y})_{x,y\in\mathcal{A}}$ of transition rates. The meaning of this matrix is that, when the state of a site is x, it is turned to y at rate $s_{x,y}$, independently of the other nucleotides. For S to be an RN matrix, some identities between coefficients must hold. Specifically, S must be of the following form:

$$\begin{array}{c} \\ A \\ T \\ C \\ G \end{array} \begin{array}{c} \begin{array}{cccc} A & T & C & G \end{array} \\ \begin{pmatrix} - & v_T & v_C & w_G \\ v_A & - & w_C & v_G \\ v_A & w_T & - & v_G \\ w_A & v_T & v_C & - \end{pmatrix} \end{array}.$$

The YpR part of the model is characterized by eight transition rates r_x^y, indexed by all the couples (x, y) in $\mathcal{A} \times \mathcal{A}$ such that x and y are not both purines nor both pyrimidines. Thus the list of available YpR rates is

$$r_A^C, \quad r_T^G, \quad r_C^A, \quad r_G^T, \quad r_G^C, \quad r_T^A, \quad r_C^G, \quad r_A^T.$$

To describe the meaning of these rates, we introduce the notations

$$Y := \{C, T\},\ \mathcal{I}_A := Y =: \mathcal{I}_G,\ R := \{A, G\},\ \mathcal{I}_C := R =: \mathcal{I}_T,$$

and

$$A^* := G, \quad G^* := A, \quad C^* := T, \quad T^* := C.$$

For every x in R and every y in $\mathcal{I}_x$, if the state of a site is x and the state of its left neighbor is y, then the transition from x to x^* occurs at rate r_x^y. Similarly, for every x in Y and every y in $\mathcal{I}_x$, if the state of a site is x and the state of its right neighbor is y, then the transition from x to x^* occurs at rate r_x^y. We refer to [3] for a thorough discussion of the properties of RN+YpR models.

In the context of this paper, we use a specification of the dynamics by transition rules which is not the simplest possible one, mathematically speaking, but which enjoys coupling properties that are crucial in the sequel. We write a corresponding list of transition rules after another definition.

Here is the list of transition rules. Recall that every rule $\mathcal{R}$ is of the form $\mathcal{R} = (c, r)$ for a rate r and a context $c = (A, \ell, s)$. Accordingly, for every symbols α and β, we write $\mathcal{R}_\alpha^\beta = (c_\alpha^\beta, r_\alpha^\beta)$ and $c_\alpha^\beta = (A_\alpha^\beta, \ell_\alpha^\beta, s_\alpha^\beta)$.

- For every x in $\mathcal{A}$, the rule $\mathcal{R}^U_x$ is defined by
$$A^U_x := \emptyset,\ \ell^U_x := \emptyset,\ s^U_x := x,\ r^U_x := \min\{v_x, w_x\}.$$
- For every x in $\mathcal{A}$, the rule $\mathcal{R}^V_x$ is defined by
$$A^V_x := \{0\},\ \ell^V_x := \mathcal{I}_x,\ s^V_x := x,\ r^V_x := (v_x - w_x)^+.$$
- For every x in $\mathcal{A}$, the rule $\mathcal{R}^W_x$ is defined by
$$A^W_x := \{0\},\ \ell^W_x := \mathcal{A} \setminus \mathcal{I}_x,\ s^W_x := x,\ r^W_x := (w_x - v_x)^+.$$
- For every $\{x, y\} = Y$ and every z in R, the rule $\mathcal{R}^Y_{xz,yz}$ is
$$A^Y_{xz,yz} := \{0, +1\},\ s^Y_{xz,yz} := y,\ r^Y_{xz,yz} := r^z_y,\ \ell^Y_{xz,yz} := \{(x, z)\}.$$
- For every $\{x, y\} = R$ and every z in Y, the rule $\mathcal{R}^R_{zx,zy}$ is defined by
$$A^R_{zx,zy} := \{-1, 0\},\ s^R_{zx,zy} := y,\ r^R_{zx,zy} := r^z_y,\ \ell^R_{zx,zy} := \{(z, x)\}.$$

In the following, the non-degeneracy condition (ND) holds:

(ND) For every nucleotide x in $\mathcal{A}$, w_x and v_x are positive.

5.1.2. *Perturbed RN+YpR models.* We consider perturbations equivalent to the addition to the RN rules described above, of a generic matrix of substitution rates, that may not satisfy the RN property, and to the addition to the YpR rules, of generic neighbor-dependent transition rates, where the dependence is either to the right neighbor or to the left neighbor.

Here is the list of perturbative transition rules.

- For every distinct x and y in $\mathcal{A}$, the rule $\mathcal{R}^\varepsilon_{x,y}$ is
$$A^\varepsilon_{x,y} := \{0\},\ \ell^\varepsilon_{x,y} := \{x\},\ s^\varepsilon_{x,y} := y,\ r^\varepsilon_{x,y} := \varepsilon(x, y).$$
- For every distinct x and y and every z in $\mathcal{A}$, the rule $\mathcal{R}^\varepsilon_{zx,zy}$ is
$$A^\varepsilon_{zx,zy} := \{-1, 0\},\ \ell^\varepsilon_{zx,zy} := \{(z, x)\},\ s^\varepsilon_{zx,zy} := y,\ r^\varepsilon_{zx,zy} := \varepsilon(zx, zy).$$
- For every distinct x and y and every z in $\mathcal{A}$, the rule $\mathcal{R}^\varepsilon_{xz,yz}$ is
$$A^\varepsilon_{xz,yz} := \{0, +1\},\ \ell^\varepsilon_{xz,yz} := \{(x, z)\},\ s^\varepsilon_{xz,yz} := y,\ r^\varepsilon_{xz,yz} := \varepsilon(xz, yz).$$

5.2. **Two coupling times with ambiguities.**

Notation 5.1. *For every site x and every subset $K =: \{\mathcal{R}_i\,;\, i \in J\}$ of $\mathfrak{R}$ with $J \subset \mathfrak{I}$, let*
$$\Psi(x, K) := \Psi(x, J) = \bigcup_{i \in J} \Psi(x, i), \quad \Psi(K) := \Psi(J),$$
and
$$r(K) := \sum_{i \in J} r_i.$$
By an abuse of notation, in the rest of the paper, we also use the shorthands $\Psi(x, \mathcal{R}_i) := \Psi(x, i)$ and $\Psi(\mathcal{R}_i) := \Psi(i)$ for every i in $\mathfrak{I}$.

Definition 5.2. *Let $\mathfrak{Z}_+$, $\mathfrak{Z}_0$, $\mathfrak{Z}_-$, $\mathfrak{Z}'_+$, $\mathfrak{Z}'_-$ and $\mathfrak{P}$ denote subsets of the rule set $\mathfrak{R}$ such that the sets $\mathfrak{Z}_+$, $\mathfrak{Z}_0$, $\mathfrak{Z}_-$ are not empty, and the sets*

$$\mathfrak{Z}_+ \cap \mathfrak{Z}'_+,\ \mathfrak{Z}_- \cap \mathfrak{Z}'_-,\ \mathfrak{Z}_0 \cap \mathfrak{P},\ \mathfrak{Z}_+ \cap \mathfrak{P},\ \mathfrak{Z}_- \cap \mathfrak{P},\ \mathfrak{Z}'_+ \cap \mathfrak{Z}'_- \cap \mathfrak{P}$$

are all empty. Let $\mathfrak{Z} := (\mathfrak{Z}_+, \mathfrak{Z}_0, \mathfrak{Z}_-, \mathfrak{Z}'_+, \mathfrak{Z}'_-)$.

Let t_-, t_0 and t_+ denote negative real numbers. Say that a coupling event based on $\mathfrak{Z}$ occurs at site $x = 0$ and at times (t_-, t_0, t_+) if the following holds.

- *$t_- < t_0$ and $t_+ < t_0$.*
- *t_- belongs to $\Psi(-1, \mathfrak{Z}_-)$, t_+ belongs to $\Psi(+1, \mathfrak{Z}_+)$ and t_0 belongs to $\Psi(0, \mathfrak{Z}_0)$.*
- *The sets $\Psi(-1, \mathfrak{Z}'_-) \cap]t_-, t_0[$ and $\Psi(+1, \mathfrak{Z}'_+) \cap]t_+, t_0[$ are both empty.*

Let $T_{\mathfrak{Z}}$ denote the maximum of the times $\min\{t_-, t_+\}$ such that a coupling event based on $\mathfrak{Z}$ occurs at times (t_-, t_0, t_+). Let $H_{\mathfrak{Z}}^{\mathfrak{P}}$ denote the set

$$H_{\mathfrak{Z}}^{\mathfrak{P}} := (\Psi(-1, \mathfrak{P}) \cup \Psi(0, \mathfrak{P}) \cup \Psi(+1, \mathfrak{P})) \cap [T_{\mathfrak{Z}}, 0[.$$

Remark 5.3. *When there exists at least one triple (t_-, t_0, t_+) which corresponds to a coupling event based on $\mathfrak{Z}$, $T_{\mathfrak{Z}}$ is indeed a maximum since the set*

$$(\Psi(-1) \cup \Psi(0) \cup \Psi(+1)) \cap]\min\{t_-, t_+\}, 0[$$

is finite. When there exists no triple (t_-, t_0, t_+) which corresponds to a coupling event based on $\mathfrak{Z}$, let $T_{\mathfrak{Z}} := -\infty$.

We define two examples of coupling events in the context of perturbed RN+YpR models, that we call YpR sensitive and YpR insensitive.

Definition 5.4. *YpR sensitive coupling events are based on the following choice of $\mathfrak{Z}$ and $\mathfrak{P}$.*

- $\mathfrak{Z}_0 := \{\mathcal{R}_y^U\ ;\ y \in \mathcal{A}\}$.
- $\mathfrak{Z}_- := \{\mathcal{R}_y^V\ ;\ y \in R\} \cup \{\mathcal{R}_y^U\ ;\ y \in \mathcal{A}\}$.
- $\mathfrak{Z}_+ := \{\mathcal{R}_y^V\ ;\ y \in Y\} \cup \{\mathcal{R}_y^U\ ;\ y \in \mathcal{A}\}$.
- $\mathfrak{Z}'_- := \{\mathcal{R}_{xz,yz}^Y\ ;\ \{x, y\} = Y,\ z \in R\}$.
- $\mathfrak{Z}'_+ := \{\mathcal{R}_{zx,zy}^R\ ;\ \{x, y\} = R,\ z \in Y\}$.
- $\mathfrak{P} := \{\mathcal{R}_{zx,zy}^\varepsilon, \mathcal{R}_{xz,yz}^\varepsilon, \mathcal{R}_{x,y}^\varepsilon\ ;\ (x, y, z) \in \mathcal{A}^3\}$.

Definition 5.5. *YpR insensitive coupling events are based on the following choice of $\mathfrak{Z}$ and $\mathfrak{P}$.*

- $\mathfrak{Z}_0 := \{\mathcal{R}_y^U\ ;\ y \in \mathcal{A}\}$.
- $\mathfrak{Z}_- := \{\mathcal{R}_y^V, \mathcal{R}_y^U\ ;\ y \in R\}$.
- $\mathfrak{Z}_+ := \{\mathcal{R}_y^V, \mathcal{R}_y^U\ ;\ y \in Y\}$.
- $\mathfrak{Z}'_- := \{\mathcal{R}_y^U, \mathcal{R}_y^V, \mathcal{R}_{x,y}^\varepsilon, \mathcal{R}_{xz,yz}^\varepsilon, \mathcal{R}_{zx,zy}^\varepsilon\ ;\ x \in \mathcal{A},\ y \in Y,\ z \in \mathcal{A}\}$.
- $\mathfrak{Z}'_+ := \{\mathcal{R}_y^U, \mathcal{R}_y^V, \mathcal{R}_{x,y}^\varepsilon, \mathcal{R}_{xz,yz}^\varepsilon, \mathcal{R}_{zx,zy}^\varepsilon\ ;\ x \in \mathcal{A},\ y \in R,\ z \in \mathcal{A}\}$.
- $\mathfrak{P} := \{\mathcal{R}_{x,y}^\varepsilon, \mathcal{R}_{xz,yz}^\varepsilon, \mathcal{R}_{zx,zy}^\varepsilon\ ;\ (x, y, z) \in \mathcal{A}^3\}$.

Notation 5.6. *We write T_{sen} and H_{sen}, respectively T_{ins} and H_{ins}, for $T_{\mathfrak{Z}}$ and $H_{\mathfrak{Z}}^{\mathfrak{P}}$ associated to a YpR sensitive, respectively YpR insensitive, coupling event.*

Here is our main result about these random variables.

Theorem E. *The random variable $(T_{\mathrm{sen}}, H_{\mathrm{sen}})$ defines a coupling time with ambiguities, whose width is bounded by $a_+ = a_- = 2$, and such that T_{sen} is exponentially integrable. The same assertions hold for $(T_{\mathrm{ins}}, H_{\mathrm{ins}})$.*

Remark 5.7. *Due to the non-degeneracy assumption (ND), and to standard independence properties of Poisson processes, T_{sen} and T_{ins} are almost surely finite negative random variables. Properties (2), (3) and (4) of definition 2.10 are also easy to establish, so the real issue is to prove property (5) and the exponential integrability.*

5.3. **Proof of the main result.** This section is devoted to the proof of theorem E.

5.3.1. *Preliminary result.*

Notation 5.8. *Let ϱ denote the application which fuses the two purines together, and η the application which fuses the two pyrimidines together, that is*

$$\varrho(A) := R =: \varrho(G), \quad \varrho(C) := C, \quad \varrho(T) := T,$$

and

$$\eta(A) := A, \quad \eta(G) := G, \quad \eta(C) := Y =: \eta(T).$$

For every times $s < t$, every configuration ξ, let

$$\phi_0(\xi, s, t, \Psi) := (\varrho(\Phi(\xi, s, t, -1)), \Phi(\xi, s, t, 0), \eta(\Phi(\xi, s, t, +1))) .$$

Lemma 5.9. *Assume that $\mathrm{Perf}(\xi, t, H_{\mathrm{sen}}) = \mathrm{Perf}(\xi', t, H_{\mathrm{sen}})$ for a time $t < T_{\mathrm{sen}}$ and for some configurations ξ and ξ'. Then, $\mathbb{P}$ almost surely,*

$$\phi_0(\xi, t, t_0, \Psi) = \phi_0(\xi', t, t_0, \Psi).$$

The same statement holds if one replaces T_{sen} and H_{sen} by T_{ins} and H_{ins}, respectively.

From now on, we assume that $\Omega_1 \cap \Omega_2 \cap \{T_{\mathrm{sen}} > -\infty\} \cap \{T_{\mathrm{ins}} > -\infty\}$ holds, since this event has full probability.

5.3.2. *Proof of the preliminary result for YpR sensitive coupling events.* We study what happens if one starts at a time $t < T_{\mathrm{sen}}$ from two initial configurations ξ and ξ', such that $\mathrm{Perf}(\xi, t, H_{\mathrm{sen}}) = \mathrm{Perf}(\xi', t, H_{\mathrm{sen}})$.

Let $t_1 < \cdots < t_r$ denote an ordering of $\Psi(-1) \cap [t_-, t_0[$.

Claim 5.10. *With full probability, for every $t < T_{\mathrm{sen}}$, for every index $1 \leqslant k \leqslant r$,*

$$\varrho(\Phi(\xi, t, t_k, -1)) = \varrho(\Phi(\xi', t, t_k, -1)).$$

If claim 5.10 holds, $\varrho(\Phi(\xi, t, t_0^-, -1)) = \varrho(\Phi(\xi', t, t_0^-, -1))$, and this fact implies that $\varrho(\Phi(\xi, t, t_0, -1)) = \varrho(\Phi(\xi', t, t_0, -1))$.

A symmetric argument shows that $\eta(\Phi(\xi, t, t_0, +1)) = \eta(\Phi(\xi', t, t_0, +1))$.

Finally, at site 0 and time t_0, the definition of T_{sen} implies that there is a rule of the form $\mathcal{R}_y^U$, hence $\Phi(\xi, t, t_0, 0) = \Phi(\xi', t, t_0, 0)$ and lemma 5.9 holds for T_{sen}.

Proof of claim 5.10. Induction on k. Start with $t_1 = t_-$. By definition of T_{sen}, at site -1 and time t_-, there is a point corresponding to a rule among $\mathcal{R}_x^U$ and $\mathcal{R}_y^V$ for any x and any purine y. Each rule $\mathcal{R}_x^U$ yields a nucleotide x for both initial conditions ξ and ξ'. As regards the rules $\mathcal{R}_A^V$ and $\mathcal{R}_G^V$, either there is a purine at site -1 and time $(t_-)^-$, in which case the rule is not performed, or there is a pyrimidine and the rule is performed. In both cases, there is a purine at site -1 and time t_-. This proves that $\varrho(\Phi(\xi, t, t_-, -1)) = \varrho(\Phi(\xi', t, t_-, -1))$, hence the claim holds for $k = 1$.

Now we assume that the claim holds for $k-1$ with $k \leqslant r$, hence $\varrho(\Phi(\xi, t, t_k^-, -1)) = \varrho(\Phi(\xi', t, t_k^-, -1))$ and we consider the effect of the rule applied at time t_k. Call this rule $\mathcal{R}$. Several cases arise.

- If $\mathcal{R}$ is perturbative, $\mathrm{Perf}(\xi, t, H_{\mathrm{sen}}) = \mathrm{Perf}(\xi', t, H_{\mathrm{sen}})$, hence $\mathcal{R}$ is performed for both initial conditions ξ and ξ', or for none. In both cases, $\varrho(\Phi(\xi, t, t_k, -1)) = \varrho(\Phi(\xi', t, t_k, -1))$.
- The same reasoning holds if $\mathcal{R}$ is non-perturbative and of the form $\mathcal{R}_y^U$.
- If $\mathcal{R} = \mathcal{R}_y^V$, since $\varrho(\Phi(\xi, t, t_k^-, -1)) = \varrho(\Phi(\xi', t, t_k^-, -1))$ by assumption, $\Phi(\xi, t, t_k^-, -1)$ and $\Phi(\xi', t, t_k^-, -1)$ are both purines or both pyrimidines. Hence, $\mathcal{R}$ is performed for both configurations ξ and ξ' or for none, and $\varrho(\Phi(\xi, t, t_k, -1)) = \varrho(\Phi(\xi', t, t_k, -1))$.
- If $\mathcal{R} = \mathcal{R}_y^W$ for a purine y, the application of $\mathcal{R}$ leaves ϱ unchanged, since $\mathcal{R}$ can only turn an A to a G or vice-versa.
- If $\mathcal{R} = \mathcal{R}_y^W$ for a pyrimidine y, $\varrho(\Phi(\xi, t, t_k^-, -1)) = \varrho(\Phi(\xi', t, t_k^-, -1))$, hence $\mathcal{R}$ is performed for both ξ and ξ', or for none.
- If $\mathcal{R} = \mathcal{R}_{zx,zy}^R$, the application of $\mathcal{R}$ leaves ϱ unchanged, since $\mathcal{R}$ can only turn an A to a G or vice-versa.
- Finally, the definition of T_{sen} excludes the rules $\mathcal{R}_{sz,yz}^Y$.

This proves claim 5.10. □

5.3.3. *Proof of the preliminary result for YpR insensitive coupling events.* As in the proof for T_{sen}, let $t_1 < \cdots < t_r$ denote an ordering of $\Psi(-1) \cap [t_-, t_0[$.

Claim 5.11. *With full probability, for every time $t < T_{\mathrm{ins}}$ and index $1 \leqslant k \leqslant r$,*

$$\varrho(\Phi(\xi, t, t_k, -1)) = \varrho(\Phi(\xi', t, t_k, -1)) = R.$$

If claim 5.11 holds, $\varrho(\Phi(\xi, t, t_0^-, -1)) = \varrho(\Phi(\xi', t, t_0^-, -1))$, and this fact implies that $\varrho(\Phi(\xi, t, t_0, -1)) = \varrho(\Phi(\xi', t, t_0, -1))$.

A symmetric argument shows that $\eta(\Phi(\xi, t, t_0, +1)) = \eta(\Phi(\xi', t, t_0, +1))$.

Finally, at site 0 and time t_0, the definition of T_{ins} implies that there is a rule of the form $\mathcal{R}_y^U$, hence $\Phi(\xi, t, t_0, 0) = \Phi(\xi', t, t_0, 0)$, and lemma 5.9 holds for T_{ins} as well.

Proof of claim 5.11. Induction on k. Start with $t_1 = t_-$. By the definition of T_{ins}, at site -1 and time t_-, there is a point corresponding to a rule $\mathcal{R}_y^U$ or $\mathcal{R}_y^V$ with a purine y. Hence $\varrho(\Phi(\xi, t, t_-, -1)) = \varrho(\Phi(\xi', t, t_-, -1))$ is a purine.

Let $k \leqslant r$ and assume that the claim holds for $k-1$, hence $\varrho(\Phi(\xi, t, t_{k-1}, -1))$ and $\varrho(\Phi(\xi', t, t_{k-1}, -1))$ coincide. Thus, $\varrho(\Phi(\xi, t, t_k^-, -1))$ and $\varrho(\Phi(\xi', t, t_k^-, -1))$

coincide. Consider the effect of the rule applied at time t_k. Call this rule $\mathcal{R}$. Several cases arise.

- Assume first that $\mathcal{R}$ is perturbative. If $\mathcal{R}$ is performed, $\mathcal{R}$ leads to a purine because, by the definition of T_{ins}, every perturbative rule applied at times $t_1, \dots, t_n$, leads to a purine. If $\mathcal{R}$ is not performed, the induction hypothesis shows that we must as well have a purine. As a consequence, whether $\mathcal{R}$ is performed or not, $\varrho(\Phi(\xi, t, t_k, -1)) = \varrho(\Phi(\xi', t, t_k, -1))$ is a purine.
- The same conclusion holds if $\mathcal{R} = \mathcal{R}_y^U$ or $\mathcal{R}_y^V$, because in this case, by the definition of T_{ins}, y must be a purine.
- If $\mathcal{R} = \mathcal{R}_y^W$ for a pyrimidine y, the induction hypothesis implies that $\mathcal{R}$ is not performed.
- If $\mathcal{R} = \mathcal{R}_y^W$ for a purine y, whether $\mathcal{R}$ is performed or not does not affect the value of ϱ.
- If $\mathcal{R} = \mathcal{R}_{zx,zy}^R$, performing $\mathcal{R}$ has no effect on the value of ϱ, since $\mathcal{R}$ can only turn an A to a G or vice-versa.
- Rules $\mathcal{R}_{xz,yz}^Y$ are not performed since $\Phi(\xi, t, t_k, -1)$ and $\Phi(\xi', t, t_k, -1)$ are both purines, by the induction hypothesis.

This proves claim 5.11. □

5.3.4. *Application to theorem E.* Let T denote T_{sen} or T_{ins}. Let $t_1 < \cdots < t_r$ denote an ordering of the set $(\Psi(-1) \cup \Psi(0) \cup \Psi(+1)) \cap]t_0, 0[$.

Claim 5.12. *With full probability, for every time $t < T$ and index $1 \leqslant k \leqslant r$,*

$$\phi_0(\xi, t, t_k, \Psi) = \phi_0(\xi', t, t_k, \Psi).$$

Claim 5.12 shows that property (5) in definition 2.10 holds. The boundedness of the width is then straightforward.

Proof of claim 5.12. Induction on k. For $k = 1$, this is lemma 5.9. Let $k \leqslant r$, assume that the claim holds for t_{k-1} and call $\mathcal{R}$ the rule applied at time t_k. Let x_k denote the corresponding site, hence t_k is in $\Psi(x_k)$.

- To begin with, if $\mathcal{R}$ is perturbative, $\mathcal{R}$ is performed for both initial conditions ξ and ξ', or for none, hence the claim holds for t_k.
- The same is true if $\mathcal{R} = \mathcal{R}_y^U$ for a given y.
- If $\mathcal{R} = \mathcal{R}_y^V$ for a given y, $\phi_0(\xi, t, t_{k-1}, \Psi) = \phi_0(\xi', t, t_{k-1}, \Psi)$, hence, for $x = -1$, 0 or $+1$, $\Phi(\xi, t, t_k^-, x)$ and $\Phi(\xi', t, t_k^-, x)$ are both purines or both pyrimidines. This means that $\mathcal{R}$ is performed for both ξ and ξ' or for none, and the claim holds for t_k.
- If $\mathcal{R} = \mathcal{R}_y^W$ for a given y, several subcases arise.
  - If $x_k = -1$ and y is a purine, performing $\mathcal{R}$ has no effect on the value of ϱ, since $\mathcal{R}$ can only turn an A to a G or vice-versa.
  - If $x_k = -1$ and y is a pyrimidine, $\mathcal{R}$ is performed for both ξ and ξ' or for none, because $\varrho(\Phi(\xi, t, t_k^-, -1)) = \varrho(\Phi(\xi', t, t_k^-, -1))$.
  - If $x_k = +1$, symmetric arguments hold.
  - If $x_k = 0$, $\Phi(\xi, t, t_k^-, -1) = \Phi(\xi', t, t_k^-, -1)$, hence $\mathcal{R}$ is performed for both ξ and ξ', or for none.

  This concludes the case when $\mathcal{R} = \mathcal{R}_y^W$ for a given y.

- If $\mathcal{R}$ is a rule $\mathcal{R}^R_{zx,zy}$ and $x_k = -1$, performing $\mathcal{R}$ has no effect on the value of ϱ, since it can only turn an A to a G or vice-versa.
- If $\mathcal{R}$ is a rule $\mathcal{R}^Y_{xz,yz}$ and $x_k = -1$, the fact that $\varrho(\Phi(\xi, t, t_k^-, -1))$ and $\varrho(\Phi(\xi', t, t_k^-, -1))$ are equal and the fact that $\Phi(\xi, t, t_k^-, 0)$ and $\Phi(\xi', t, t_k^-, 0)$ are equal ensures that $\mathcal{R}$ is performed for both ξ and ξ', or for none.
- If $\mathcal{R}$ is a rule $\mathcal{R}^R_{zx,zy}$ or a rule $\mathcal{R}^Y_{xz,yz}$, and $x_k = +1$, similar arguments hold.
- If $x_k = 0$, the facts that $\varrho(\Phi(\xi, t, t_k^-, -1))$ and $\varrho(\Phi(\xi', t, t_k^-, -1))$ are equal, that $\Phi(\xi, t, t_k^-, 0)$ and $\Phi(\xi', t, t_k^-, 0)$ are equal for a rule of the form $\mathcal{R}^R_{zx,zy}$, and the facts that $\eta(\Phi(\xi, t, t_k^-, +1))$ and $\eta(\Phi(\xi', t, t_k^-, +1))$ are equal and that $\Phi(\xi, t, t_k^-, 0)$ and $\Phi(\xi', t, t_k^-, 0)$ are equal for a rule of the form $\mathcal{R}^Y_{xz,yz}$ ensure that $\mathcal{R}$ is performed for both ξ and ξ', or for none.

This concludes the proof of claim 5.12. □

5.3.5. *End of the proof of theorem E.* To conclude the proof of theorem E, one must show that T_{sen} and T_{ins} are both exponentially integrable. The proof is the same in both cases.

We define inductively decreasing sequences of random times $(U_n)_{n\geqslant 0}$, $(U_n^-)_{n\geqslant 0}$, $(U_n^+)_{n\geqslant 0}$ and $(U_n^0)_{n\geqslant 0}$. Let $U_0 := 0$. For every $n \geqslant 0$, let

$$U_n^- := \max \Psi(-1, \mathfrak{Z}_-) \cap]-\infty, U_n[, \quad U_n^+ := \max \Psi(+1, \mathfrak{Z}_+) \cap]-\infty, U_n[,$$

and

$$U_n^0 := \min\{U_n^-, U_n^+\}, \quad U_{n+1} := \max \Psi(0, \mathfrak{Z}_0) \cap]-\infty, U_n^0[.$$

Hence $U_{n+1} < U_n^0 \leqslant U_n^-, U_n^+ < U_n$ and every U_n is almost surely finite.

For every $n \geqslant 0$, consider the event

$$S_n := \left\{ \Psi(-1, \mathfrak{Z}'_-) \cap]U_n^-, U_n[= \Psi(+1, \mathfrak{Z}'_+) \cap]U_n^+, U_n[= \emptyset \right\}.$$

For every $n \geqslant 1$, on the event S_n, the choice $t_0 = U_n$, $t_- = U_n^-$, $t_+ = U_n^+$ yields an admissible triple (t_-, t_0, t_+), hence $T_{\mathfrak{Z}} \geqslant U_{n+1}$. Let $T := U_{N+1}$ where N is the first integer $n \geqslant 1$ such that S_n holds. We wish to show that T is exponentially integrable.

The sequence $(S_n, U_n - U_{n+1})_{n\geqslant 0}$ is i.i.d. and, for every $n \geqslant 1$,

$$\{N = n\} = S_1^c \cap \cdots \cap S_{n-1}^c \cap S_n.$$

Hence, for every real number λ,

$$\Lambda_T(\lambda) = \sum_{n\geqslant 1} \mathbb{E}[\mathrm{e}^{-\lambda U_{n+1}} \,;\, N = n] = \mathbb{E}[\mathrm{e}^{-\lambda U_1}] \mathbb{E}[\mathrm{e}^{-\lambda U_1} \,;\, S_0] \sum_{n\geqslant 1} \mathbb{E}[\mathrm{e}^{-\lambda U_1} \,;\, S_0^c]^{n-1}.$$

This shows that $\Lambda_T(\lambda)$ is finite if and only if $\mathbb{E}[\mathrm{e}^{-\lambda U_1}]$ is finite and $\mathbb{E}[\mathrm{e}^{-\lambda U_1} \,;\, S_0^c] < 1$ (but we recall that T is not $T_{\mathfrak{Z}}$).

By construction, $U_1 \geqslant U_0^- + U_0^+ + (U_1 - U_0^0)$ and these three random variables are independent and exponential with parameters $r(\mathfrak{Z}_-)$, $r(\mathfrak{Z}_+)$ and $r(\mathfrak{Z}_0)$, respectively, hence $\mathbb{E}[\mathrm{e}^{-\lambda U_1}]$ is finite for every λ smaller than these three rates. Since $U_0^0 \geqslant U_1$, the same conclusion applies to U_0^0.

Furthermore, $U_0^0 - U_1$ is independent of (S_0, U_0^0) and its distribution is exponential of parameter $r(\mathfrak{Z}_0)$. Conditionally on U_0^+ and U_-^0, the number of points in

the sets $\Psi(-1,\mathfrak{Z}'_-)\cap]U_0^-,0[$ and $\Psi(+1,\mathfrak{Z}'_+)\cap]U_0^+,0[$ are independent and Poisson of parameters $-r(\mathfrak{Z}'_-)U_0^-$ and $-r(\mathfrak{Z}'_+)U_0^+$. For every $\lambda < r(\mathfrak{Z}_0)$, this yields

$$\mathbb{E}[\mathrm{e}^{-\lambda U_1}\,;\,S_0^c] = \frac{r(\mathfrak{Z}_0)}{r(\mathfrak{Z}_0)-\lambda}\,\mathbb{E}\left[(1-\mathrm{e}^{r(\mathfrak{Z}'_-)U_0^- + r(\mathfrak{Z}'_+)U_0^+})\,\mathrm{e}^{-\lambda U_0^0}\right].$$

Hence the fact that $\Lambda_T(\lambda)$ is finite for some positive λ follows from the claim below, with $r_0 := r(\mathfrak{Z}_0)$, $W := 1-\mathrm{e}^{r(\mathfrak{Z}'_-)U_0^- + r(\mathfrak{Z}'_+)U_0^+}$ and $V := -U_0^0$.

Claim 5.13. *Let V and W denote positive random variables and r_0 a positive real number. Assume that V is exponentially integrable and that $W < 1$ almost surely. For every real number $\lambda < r_0$, let*

$$F(\lambda) := \frac{r_0}{r_0-\lambda}\,\mathbb{E}[W\mathrm{e}^{\lambda V}].$$

Then there exists some positive values of λ such that $F(\lambda) < 1$.

Proof of claim 5.13. Expansions of the exponentials at order 1 with respect to the parameter λ yield $F(0) = \mathbb{E}[W] < 1$, and $F'(0) = \mathbb{E}[WV] + r_0^{-1}\mathbb{E}[W]$. Hence $F'(0)$ is finite and the proof of the claim is complete. □

5.4. **Remarks.** First, in our two examples, properties (1), (2), (3) and (5) in definition 2.10 still hold if one removes the points in $\Psi(0,\mathfrak{P})\cap]T_{\mathfrak{Z}},t_0[$ from the definition of $H_{\mathfrak{Z}}^{\mathfrak{P}}$ (here (t_-,t_0,t_+) denotes a triple such that there exists a coupling event based on $\mathfrak{Z}$ at time (t_-,t_0,t_+) and such that $\min\{t_-,t_+\}$ is maximal among such coupling events). This leads to a smaller growth parameter and Laplace transform, but, on the other hand, property (4) (the stopping property of $H_{\mathfrak{Z}}^{\mathfrak{P}}$) is lost, and we need property (4) to prove effective estimates on the ambiguity process.

Second, as regards YpR sensitive coupling events, T_{sen} is defined purely in terms of non-perturbative rules, and, in fact, the proof of theorem E implies that T_{sen} is an ordinary coupling time for the unperturbed dynamics. In other words, the content of theorem E is that the fact of fixing the ambiguities associated to some perturbative rules in $\Psi\cap]T_{\mathrm{sen}},0[$ restores the coupling property of T_{sen}.

Third, the situation is a bit different for YpR insensitive coupling events, whose definition involves both perturbative and non-perturbative rules. Indeed, removing the perturbative rules from the definitions of $\mathfrak{Z}'_-$ and $\mathfrak{Z}'_+$ for YpR insensitive coupling events would make T_{ins} a coupling time for the unperturbed dynamics. However, fixing ambiguities associated to perturbative rules in $\Psi\cap]T_{\mathrm{ins}},0[$ is not enough to restore the coupling property of T_{ins}. This is the reason why, in this example, one must introduce perturbative rules in the definition of T_{ins}.

Finally, note that it is possible to use our two examples of coupling times with ambiguities to perform perfect simulation according to the Propp-Wilson method (see [14]). Indeed, by the definition of a coupling time with ambiguities, for any finite subset $B \subset \mathbb{Z}$, the random times T_B^* are coalescence times that allow to sample exactly from the projection π_B of the invariant distribution of the particle system onto the sites in B. Moreover, from their definition, our two examples allow for an efficient detection of coalescence by an algorithm.

5.5. **Computations of growth parameters.** We recall notation 5.1 in section 5.2 and we define inductively sequences $(\lambda_n)_{n\geqslant 0}$, $(\kappa_n)_{n\geqslant 1}$ and $(\chi_n)_{n\geqslant 1}$ of times and a sequence $(\beta_n)_{n\geqslant 1}$ of bits, as follows. Let $\lambda_0 := 0$. For every integer $n \geqslant 1$, let

$$\kappa_n := \sup \Psi(0, \mathfrak{Z}_0) \cap]-\infty, \lambda_{n-1}[,$$

and

$$\chi_n := \sup(\Psi(+1, \mathfrak{Z}_+ \cup \mathfrak{Z}'_+) \cup \Psi(-1, \mathfrak{Z}_- \cup \mathfrak{Z}'_-)) \cap]-\infty, \kappa_n[.$$

We define λ_n and β_n as follows.

- If χ_n is in $\Psi(+1, \mathfrak{Z}_+)$, let $\lambda_n := \sup \Psi(-1, \mathfrak{Z}_- \cup \mathfrak{Z}'_-) \cap]-\infty, \kappa_n[$; then, if λ_n is in $\Psi(-1, \mathfrak{Z}_-)$, let $\beta_n := 1$, otherwise let $\beta_n := 0$.
- If χ_n is in $\Psi(-1, \mathfrak{Z}_-)$, let $\lambda_n := \sup \Psi(+1, \mathfrak{Z}_+ \cup \mathfrak{Z}'_+) \cap]-\infty, \kappa_n[$; then, if λ_n is in $\Psi(+1, \mathfrak{Z}_+)$, let $\beta_n := 1$, otherwise let $\beta_n := 0$.
- If χ_n is in $\Psi(+1, \mathfrak{Z}'_+)$ or in $\Psi(-1, \mathfrak{Z}'_-)$, let $\lambda_n := \chi_n$ and $\beta_n := 0$.

For every rule $\mathcal{R}$ in $\mathfrak{P}$ and $n \geqslant 1$, let

$$N_n(\mathcal{R}) := \#(\Psi(-1, \mathcal{R}) \cup \Psi(0, \mathcal{R}) \cup \Psi(+1, \mathcal{R})) \cap [\lambda_n, \lambda_{n-1}[.$$

By the independence properties of Poisson processes, for every rule $\mathcal{R}$ in $\mathfrak{P}$, the sequence $(\beta_n, N_n(\mathcal{R}))_{n\geqslant 1}$ is i.i.d. Let

$$G(\mathfrak{Z}) := \min\{n \geqslant 1\,;\, \beta_n = 1\}.$$

The distribution of $G(\mathfrak{Z})$ is geometric on $\{1, 2, \ldots\}$ and non degenerate under our non-degeneracy assumption that every r_y^U is positive.

The proof of the following lemma is an easy consequence of the definitions above and is omitted.

Lemma 5.14. $\mathbb{P}[T_{\mathfrak{Z}} = \lambda_{G(\mathfrak{Z})}] = 1$.

As a consequence, with probability one, for every rule $\mathcal{R}$ in $\mathfrak{P}$,

$$[T_{\mathfrak{Z}}, 0[= \bigcup_{n=1}^{G(\mathfrak{Z})} [\lambda_n, \lambda_{n-1}[,$$

so that

$$\#H_{\mathfrak{Z}}^{\mathfrak{P}} \cap \Psi(\mathcal{R}) = \sum_{n=1}^{G(\mathfrak{Z})} N_n(\mathcal{R}),$$

and

$$\mathbb{E}[\#H_{\mathfrak{Z}}^{\mathfrak{P}} \cap \Psi(\mathcal{R})] = \mathbb{P}[\beta_1 = 1]^{-1} \mathbb{E}[N_1(\mathcal{R})].$$

The proposition below follows.

Proposition 5.15. *The mean of* $(T_{\mathfrak{Z}}, H_{\mathfrak{Z}}^{\mathfrak{P}})$ *is*

$$m(T_{\mathfrak{Z}}, H_{\mathfrak{Z}}^{\mathfrak{P}}) = \mathbb{P}[\beta_1 = 1]^{-1} \sum_{\mathcal{R}_i \in \mathfrak{P}} (\#A_i)\, \mathbb{E}[N_1(\mathcal{R}_i)].$$

Section 5.6 describes a tree of successive conditional distributions leading, for every rule $\mathcal{R}$ in $\mathfrak{P}$, to an explicit representation of the joint distribution of $N_1(\mathcal{R})$ and β_1, and to the computation of $\mathbb{E}[N_1(\mathcal{R})]$ and $\mathbb{P}[\beta_1 = 1]$ in terms of rates.

Notation 5.16. *Introduce the shorthands* $r_0 := r(\mathfrak{Z}_0)$,

$$r_1 := r(\mathfrak{Z}_+),\ r_2 := r(\mathfrak{Z}'_+),\ r_3 := r(\mathfrak{Z}_+ \cup \mathfrak{Z}'_+) = r_1 + r_2,$$

and

$$r_4 := r(\mathfrak{Z}_-),\ r_5 := r(\mathfrak{Z}'_-),\ r_6 := r(\mathfrak{Z}_- \cup \mathfrak{Z}'_-) = r_4 + r_5.$$

Section 5.6 below leads to the following values.

- $\mathbb{P}[\beta_1 = 1] = \dfrac{r_1 r_4}{r_3 r_6}$.
- For every $\mathcal{R} = (c, r)$ in $\mathfrak{P}$ and not in $\mathfrak{Z}'_+ \cup \mathfrak{Z}'_-$,

$$\mathbb{E}[N_1(\mathcal{R})] = 3r\left(\frac{1}{r_0} + \frac{1}{r_3 + r_6}\left(1 + \frac{r_1}{r_6} + \frac{r_4}{r_3}\right)\right).$$

- For every $\mathcal{R} = (c, r)$ in $\mathfrak{P}$ and in $\mathfrak{Z}'_+ \cup \mathfrak{Z}'_-$,

$$\mathbb{E}[N_1(\mathcal{R})] = r\left(\frac{3}{r_0} + \frac{1}{r_3 + r_6}\left(2 + 3\frac{r_1}{r_6} + 3\frac{r_4}{r_3}\right)\right).$$

The fact that every $\mathbb{E}[N_1(\mathcal{R})]$ is bounded by the expression for the case when $\mathcal{R}$ is not in $\mathfrak{Z}'_+ \cup \mathfrak{Z}'_-$ yields the following result.

Proposition 5.17. *One has* $m(T_{\mathfrak{Z}}, H_{\mathfrak{Z}}^{\mathfrak{P}}) \leqslant 3M(\mathfrak{Z}) \sum_{\mathcal{R}_i \in \mathfrak{P}} r_i \, \#A_i$, *where*

$$M(\mathfrak{Z}) := \frac{r_3 r_6}{r_1 r_4}\left(\frac{1}{r_0} + \frac{1}{r_3 + r_6}\left(1 + \frac{r_1}{r_6} + \frac{r_4}{r_3}\right)\right).$$

5.6. **Tree decompositions of conditional distributions.** In this section, we describe the distributions of a family of random variables by placing each of them at the vertices of a tree. The distribution of the random variable placed at vertex $x.$ is conditional to the random variables placed at vertices which are ancestors of $x.$ in the tree. We use the following labelling: 1. is the root of the tree, and the children of a vertex $x.$ are $x.y.$ with $y = 1$, $y = 2$, and so on.

To avoid cumbersome notations, we denote by $(n).$ a piece of "trunk" of length n, that is, a vertex $1.1.\cdots.1.$ with n ones such that no ramification starts before it. For instance, 1.1.1.=(3)., 1.1.1.1.1.1.=(5).1. and 1.1.1.1.1.2.=(5).2. Finally, we recall the shorthands of notation 5.1 in section 5.2.

Three distinct situations arise: the rule $\mathcal{R}$ belongs to $\mathfrak{Z}'_+$ or to $\mathfrak{Z}'_-$ or to none of these two sets.

5.6.1. *First case: perturbative rules not in* $\mathfrak{Z}'_+ \cup \mathfrak{Z}'_-$. Assume that the rule $\mathcal{R} = (c, r)$ is in $\mathfrak{P}$ and not in $\mathfrak{Z}'_+ \cup \mathfrak{Z}'_-$.

Vertex 1. The distribution of $-\kappa_1$ is exponential with parameter r_0.

Vertex (2). The distribution of $\#(\Psi(-1, \mathcal{R}) \cup \Psi(0, \mathcal{R}) \cup \Psi(+1, \mathcal{R})) \cap]\kappa_1, 0[$ is Poisson with parameter $3r\,(-\kappa_1)$.

Vertex (3). The distribution of $\kappa_1 - \chi_1$ is exponential with parameter $r_3 + r_6$.

Vertex (4). The distribution of $\#(\Psi(-1, \mathcal{R}) \cup \Psi(0, \mathcal{R}) \cup \Psi(+1, \mathcal{R})) \cap]\chi_1, \kappa_1[$ is Poisson with parameter $3r\,(\kappa_1 - \chi_1)$.

Vertex (5). The probability that χ_1 belongs to $\Psi(+1, \mathfrak{Z}_+ \cup \mathfrak{Z}'_+)$ is $r_3/(r_3 + r_6)$, hence the probability that χ_1 belongs to $\Psi(-1, \mathfrak{Z}_- \cup \mathfrak{Z}'_-)$ is $r_6/(r_3 + r_6)$.

Vertex (5).1. If χ_1 belongs to $\Psi(+1, \mathfrak{Z}_+ \cup \mathfrak{Z}'_+)$, then the probability that χ_1 belongs to $\Psi(+1, \mathfrak{Z}_+)$ is r_1/r_3, hence the probability that χ_1 belongs to $\Psi(+1, \mathfrak{Z}'_+)$ is r_2/r_3.

Vertex (5).1.1. If χ_1 belongs to $\Psi(+1, \mathfrak{Z}'_+)$, then $\beta_1 = 0$ and $\lambda_1 = \chi_1$.

Vertex (5).1.2. If χ_1 belongs to $\Psi(+1, \mathfrak{Z}_+)$, then the distribution of $\chi_1 - \lambda_1$ is exponential with parameter r_6.

Vertex (5).1.2.1. The distribution of $\#(\Psi(-1, \mathcal{R}) \cup \Psi(0, \mathcal{R}) \cup \Psi(+1, \mathcal{R})) \cap]\lambda_1, \chi_1[$ is Poisson with parameter $3r\,(\chi_1 - \lambda_1)$.

Vertex (5).1.2.1.1. The probability that $\beta_1 = 1$ is r_4/r_6, hence the probability that $\beta_1 = 0$ is r_5/r_6.

Vertex (5).2. If χ_1 belongs to $\Psi(-1, \mathfrak{Z}_- \cup \mathfrak{Z}'_-)$, then the probability that χ_1 belongs to $\Psi(-1, \mathfrak{Z}_-)$ is r_4/r_6, hence the probability that χ_1 belongs to $\Psi(-1, \mathfrak{Z}'_-)$ is r_5/r_6.

Vertex (5).2.1. If χ_1 belongs to $\Psi(-1, \mathfrak{Z}'_-)$, then $\beta_1 = 0$ and $\lambda_1 = \chi_1$.

Vertex (5).2.2. If χ_1 belongs to $\Psi(-1, \mathfrak{Z}_-)$, then the distribution of $\chi_1 - \lambda_1$ is exponential with parameter r_3.

Vertex (5).2.2.1. The distribution of $\#(\Psi(-1, \mathcal{R}) \cup \Psi(0, \mathcal{R}) \cup \Psi(+1, \mathcal{R})) \cap]\lambda_1, \chi_1[$ is Poisson with parameter $3r\,(\chi_1 - \lambda_1)$.

Vertex (5).2.2.1.1. The probability that $\beta_1 = 1$ is r_1/r_3, hence the probability that $\beta_1 = 0$ is r_2/r_3.

5.6.2. *Second case: perturbative rules in* $\mathfrak{Z}'_+$. Assume that the rule $\mathcal{R} = (c, r)$ is in $\mathfrak{P} \cap \mathfrak{Z}'_+$.

Vertex 1. The distribution of $-\kappa_1$ is exponential with parameter r_0.

Vertex (2). The distribution of $\#(\Psi(-1, \mathcal{R}) \cup \Psi(0, \mathcal{R}) \cup \Psi(+1, \mathcal{R})) \cap]\kappa_1, 0[$ is Poisson with parameter $3r\,(-\kappa_1)$.

Vertex (3). The distribution of $\kappa_1 - \chi_1$ is exponential with parameter $r_3 + r_6$.

Vertex (4). The distribution of $\#(\Psi(-1, \mathcal{R}) \cup \Psi(0, \mathcal{R}) \cup \Psi(+1, \mathcal{R})) \cap]\chi_1, \kappa_1[$ is Poisson with parameter $r\,(\kappa_1 - \chi_1)$.

Vertex (5). The probability that χ_1 belongs to $\Psi(+1, \mathfrak{Z}_+ \cup \mathfrak{Z}'_+)$ is $r_3/(r_3 + r_6)$, hence the probability that χ_1 belongs to $\Psi(-1, \mathfrak{Z}_- \cup \mathfrak{Z}'_-)$ is $r_6/(r_3 + r_6)$.

Vertex (5).1. If χ_1 belongs to $\Psi(+1, \mathfrak{Z}_+ \cup \mathfrak{Z}'_+)$, then the probability that χ_1 belongs to $\Psi(+1, \mathfrak{Z}_+)$ is r_1/r_3, hence the probability that χ_1 belongs to $\Psi(+1, \mathfrak{Z}'_+)$ is r_2/r_3.

Vertex (5).1.1. If χ_1 belongs to $\Psi(+1, \mathfrak{Z}'_+)$, then $\beta_1 = 0$ and $\lambda_1 = \chi_1$, and the probability that χ_1 belongs to $\Psi(+1, \mathcal{R})$ is r/r_2.

Vertex (5).1.2. If χ_1 belongs to $\Psi(+1, \mathfrak{Z}_+)$, then the distribution of $\chi_1 - \lambda_1$ is exponential with parameter r_6.

Vertex (5).1.2.1. The distribution of $\#(\Psi(-1, \mathcal{R}) \cup \Psi(0, \mathcal{R}) \cup \Psi(+1, \mathcal{R})) \cap]\lambda_1, \chi_1[$ is Poisson with parameter $3r\,(\chi_1 - \lambda_1)$.

Vertex (5).1.2.1.1. The probability that $\beta_1 = 1$ is r_4/r_6, hence the probability that $\beta_1 = 0$ is r_5/r_6.

Vertex (5).2. If χ_1 belongs to $\Psi(-1, \mathfrak{Z}_- \cup \mathfrak{Z}'_-)$, the probability that χ_1 belongs to $\Psi(-1, \mathfrak{Z}_-)$ is r_4/r_6, hence the probability that χ_1 belongs to $\Psi(-1, \mathfrak{Z}'_-)$ is r_5/r_6.

Vertex (5).2.1. If χ_1 belongs to $\Psi(-1, \mathfrak{Z}'_-)$, then $\beta_1 = 0$ and $\lambda_1 = \chi_1$.

Vertex (5).2.2. If χ_1 belongs to $\Psi(-1, \mathfrak{Z}_-)$, then the distribution of $\chi_1 - \lambda_1$ is exponential with parameter r_3.

Vertex (5).2.2.1. The distribution of $\#(\Psi(-1, \mathcal{R}) \cup \Psi(0, \mathcal{R}) \cup \Psi(+1, \mathcal{R})) \cap]\lambda_1, \chi_1[$ is Poisson with parameter $2r\,(\chi_1 - \lambda_1)$.

Vertex (5).2.2.1.1. The probability that $\beta_1 = 1$ is r_1/r_3, hence the probability that $\beta_1 = 0$ is r_2/r_3. If $\beta_1 = 0$, the probability that χ_1 belongs to $\Psi(+1, \mathcal{R})$ is r/r_2.

5.6.3. *Third case: perturbative rules in* $\mathfrak{Z}'_-$. One can deduce this case from the second case: a similar algorithm holds, obtained through the transformations $\mathfrak{Z}_- \leftrightarrow \mathfrak{Z}_+$ and $\mathfrak{Z}'_- \leftrightarrow \mathfrak{Z}'_+$ in the steps of the algorithm for rules in $\mathfrak{Z}'_+$.

5.7. **Example: perturbed Jukes-Cantor model with CpG influence.** We now apply this upper bound to perturbations of Jukes-Cantor model of evolution with influence of the dinucleotide CpG. Namely, we assume that CpG mutates to CpA and to TpG, both at rates δ, that every nucleotide x mutates to $y \neq x$ at rate $1 + \varepsilon(x, y)$ with $\varepsilon(x, y) \geqslant 0$. Let $|\varepsilon|$ denote the sum over every x and y of the perturbations $\varepsilon(x, y)$.

As regards sensitive coupling events, $r_0 = r_1 = r_4 = 4$ and $r_2 = r_5 = \delta$. The mean of the coupling mechanism based on $(T_{\mathrm{sen}}, H_{\mathrm{sen}})$ is bounded by $\frac{3}{64}|\varepsilon|(40+10\delta+\delta^2)$, hence this coupling mechanism is subcritical as soon as $|\varepsilon| < \varepsilon_{\mathrm{sen}}(\delta)$, with

$$\varepsilon_{\mathrm{sen}}(\delta) := \frac{64}{3(40 + 10\delta + \delta^2)}.$$

Hence, $\varepsilon_{\mathrm{sen}}(0) = \frac{8}{15}$, $\varepsilon_{\mathrm{sen}}(2) = \frac{1}{3}$, $\varepsilon_{\mathrm{sen}}(10) = \frac{4}{45}$, and $\varepsilon_{\mathrm{sen}}(\delta) \to 0$ when $\delta \to +\infty$.

As regards insensitive coupling events, $r_0 = 4$, $r_1 = r_4 = 2$, r_2 is the sum of the modifications $\varepsilon(x, y)$ for y in R, and r_5 is the sum of the modifications $\varepsilon(x, y)$ for y in Y. The mean of the coupling mechanism based on $(T_{\mathrm{ins}}, H_{\mathrm{ins}})$ is bounded by $\frac{3}{64}|\varepsilon|(64 + 12|\varepsilon| + |\varepsilon|^2)$, hence this coupling mechanism is subcritical as soon as $|\varepsilon| < \varepsilon_{\mathrm{ins}}$, where $\varepsilon_{\mathrm{ins}}$ is the unique positive root of

$$3\varepsilon_{\mathrm{ins}}(64 + 12\varepsilon_{\mathrm{ins}} + \varepsilon_{\mathrm{ins}}^2) = 64.$$

For every value of δ, this insensitive coupling mechanism is subcritical as soon as $|\varepsilon| \leqslant \frac{3}{10}$.

Theorem F. *Every perturbation by rates ε of the Jukes-Cantor model with CpG influence of magnitude δ is ergodic as soon as $|\varepsilon| < \max\{\varepsilon_{\mathrm{ins}}, \varepsilon_{\mathrm{sen}}(\delta)\}$. Furthermore, every finite marginal converges exponentially fast to the corresponding finite marginal of the stationary distribution, and the correlations of the stationary distribution decay exponentially fast.*

References

[1] Eitan Altman, Konstantin E. Avrachenkov, and Rudesindo Núñez-Queija. Perturbation analysis for denumerable Markov chains with application to queueing models. *Advances in Applied Probability*, 36(3):839–853, 2004.

[2] Peter Arndt, Chris Burge, and Terence Hwa. DNA sequence evolution with neighbour-dependent mutation. *Journal of Computational Biology*, 10:313–322, 2003.
[3] Jean Bérard, Jean-Baptiste Gouéré, and Didier Piau. Solvable models of neighbor-dependent nucleotide substitution processes. *Mathematical Biosciences*, To appear.
[4] Ole Christensen, Asger Hobolth, and Jens Ledet Jensen. Pseudo-likelihood analysis of codon substitution models with neighbor dependent rates. *Journal of Computational Biology*, 12:1166–1182, 2005.
[5] Laurent Duret and Nicolas Galtier. The covariation between TpA deficiency, CpG deficiency, and G+C content of human isochores is due to a mathematical artifact. *Molecular Biology and Evolution*, 27:1620–1625, 2000.
[6] Peter W. Glynn and Sean P. Meyn. A Liapounov bound for solutions of the Poisson equation. *Ann. Probab.*, 24(2):916–931, 1996.
[7] Dick G. Hwang and Phil Green. Bayesian Markov chain Monte Carlo sequence analysis reveals varying neutral substitution patterns in mammalian evolution. *Proceedings of the National Academy of Sciences USA*, 101:13994–14001, 2004.
[8] Jens Ledet Jensen and Anne-Mette Krabbe Pedersen. Probabilistic models of DNA sequence evolution with context dependent rates of substitution. *Advances in Applied Probability*, 32(2):499–517, 2000.
[9] Tomasz Komorowski and Stefano Olla. On mobility and Einstein relation for tracers in time-mixing random environments. *Journal of Statistical Physics*, 118(3-4):407–435, 2005.
[10] Thomas M. Liggett. *Interacting particle systems*, volume 276 of *Grundlehren der Mathematischen Wissenschaften*. Springer-Verlag, New York, 1985.
[11] Thomas M. Liggett. *Stochastic interacting systems: contact, voter and exclusion processes*, volume 324 of *Grundlehren der Mathematischen Wissenschaften*. Springer-Verlag, Berlin, 1999.
[12] Gerton A. Lunter and Jotun Hein. A nucleotide substitution model with nearest-neighbour interactions. *Bioinformatics*, 20:i216–i223, 2004.
[13] Christian Maes and Karel Netočný. Spacetime expansions for weakly coupled interacting particle systems. *Journal of Physics A, Mathematical and General*, 35(13):3053–3077, 2002.
[14] James Gary Propp and David Bruce Wilson. Exact sampling with coupled Markov chains and applications to statistical mechanics. *Random Structures Algorithms*, 9(1-2):223–252, 1996.
[15] Gareth O. Roberts, Jeffrey S. Rosenthal, and Peter O. Schwartz. Convergence properties of perturbed Markov chains. *J. Appl. Probab.*, 35(1):1–11, 1998.
[16] Tony Shardlow and Andrew M. Stuart. A perturbation theory for ergodic Markov chains and application to numerical approximations. *SIAM Journal on Numerical Analysis*, 37(4):1120–1137 (electronic), 2000.
[17] Adam Siepel and David Haussler. Phylogenetic estimation of context-dependent substitution rates by maximum likelihood. *Molecular Biology and Evolution*, 21:468–488, 2004.

(Jean Bérard) Université de Lyon ; Université Lyon 1 ; Institut Camille Jordan CNRS UMR 5208 ; 43, boulevard du 11 novembre 1918, F-69622 Villeurbanne Cedex; France.
e-mail: Jean.Berard@univ-lyon1.fr

(Didier Piau) Institut Fourier - UMR 5582, Université Joseph Fourier Grenoble 1, 100 rue des Maths, BP 74, 38402 Saint Martin d'Hères, France.
e-mail: Didier.Piau@ujf-grenoble.fr